SEQ ID NO: 1
1 GGATCCCAGGGAACGTGACC ATG GTC GTA GGG ATG ACT TGA CAGTTTCAACGGGGTGCGACCACCGTTGCGC 72
1   (M   V   V   G   M   T) *   7
SEQ ID NO: 2   SEQ ID NO: 3

73 TCAGAAGGCATACGTTGGTGGAACACGTCGGAAAGCTGGGAGGTGAATCTG ATG GCT GGC GAC CAA GAG CTG 144
1   (M   A   G   D   Q   E   L   7
                                                       SEQ. ID NO: 4

145 GAA CTG CGG TTC GAC GTT CCT CTT TAC ACG CTT GCC GAG GCA TCG CGG TAC CTG GTG GTT 204
8   E   L   R   F   D   V   P   L   Y   T   L   A   E   A   S   R   Y   L   V   V   27

205 CCC CGC GCC ACC CTG GCT ACG TGG GCT GAC GGC TAC GAG CGT CGG CCG GCC AAC GCA CCG 264
28 P   R   A   T   L   A   T   W   A   D   G   Y   E   R   R   P   A   N   A   P   47

265 GCG GTC CAG GGG CAA CCG ATC GCC TTT GAC GCC TAT TCG GTC GCG CAG CTT TTT GGC GAC 324
48 A   V   Q   G   Q   P   I   A   F   D   A   Y   S   V   A   Q   L   F   G   D   67

325 GTC ACT GGT GCC CGC GTT GCG GGC GTC CAG CCG CAG CGA CAC CAC ATA CGG CCG GTC CGG 384
68 V   T   G   A   R   V   A   G   V   Q   P   Q   R   H   H   I   R   P   V   R   87

385 TTG CGG GGG CCG TTG GGT GGG GTT GGG TGC CTC CGT CAC CCC AGG CAG TTC GCT GGC TAT 444
88 L   R   G   P   L   G   G   V   G   C   L   R   H   P   R   Q   F   A   G   Y   107

445 TTG TCG CAG TAG CGCGACGGCATTGTCG ATG TCT TGG TAG CTAGCATCCGGTCGGGGGGCCGCTACCAGCG 515
108 L   S   Q) *   M   S   W   *   4

516 CCAGCGCCGGGGCTCCCCGGTCCGGGTAGTGCGCGTCGAGTTGGTCGTGGACCAGCA ATG ACT GCG ACC CGG 587
1   SEQ. ID NO: 5   (M   T   A   T   R   5

588 CGA CTT CGA AAC CGC CAC CGG TTA GAT TCC CCG ACT GCG TCA TCG CCA GGT AAA CCG CCG 647
6   R   L   R   N   R   H   R   L   D   S   P   T   A   S   S   P   G   K   P   P   25

648 GCA CTA ACG CCA GCA ACC AAC CCG TGA AGACCAACCAACGGCACCTGCGCAGGTTGCGGCTCAACCGCATC 718
26 A   L   T   P   A   T   N   P) *   34

719 ATG AAC TGC TGG ATT TCG GAC TCC CCG TAC TCT CGC GCA GTG CGT GCC CGC GAG CCT ACC 778
1(M N   C   W   I   S   D   S   P   Y   S   R   A   V   R   A   R   E   P   T   20
SEQ ID NO: 6

779 GAA GAT CGC GTG CAT GCG TTC GGC GTG GAC CGC ACA GCA CCT GGA GTT GGC GGC GCC GAG 838
21 E   D   R   V   H   A   F   G   V   D   R   T   A   P   G   V   G   G   A   E   40

839 GGC CGA GAT GGC AGG ATG ACG GAT CGT CGG GGG CGG GAA CTC CCA GGC CGC CGG ACC GTC 898
41 G   R   D   G   R   M   T   D   R   R   G   R   E   L   P   G   R   R   T   V   60

899 GCA AAC CCG TCG CAA ACC CGT CGC AAA CCG TAA GGAGTCATCC ATG AAG ACA GGC ACC GCG 959
61 A   N   P   S   Q   T   R   R   K   P) *   (M   K   T   G   T   A   6
                                                            SEQ. ID NO: 7

960 ACG ACG CGG CGC AGG CTG TTG GCA GTA CTG ATC GCC CTC GCG TTG CCG GGG GCC GCC GTT 1019
7   T   T   R   R   R   L   L   A   V   L   I   A   L   A   L   P   G   A   A   V   26

1020 GCG CTG CTG GCC GAA CCA TCA GCG ACC GGC GCG TCG GAC CCG TGC GCG GCC AGC GAA GTG 1079
27   A   L   L   A   E   P   S   A   T   G   A   S   D   P   C   A   A   S   E   V   46

1080 GCG AGG ACG GTC GGT TCG GTC GCC AAG TCG ATG GGC GAC TAC CTG GAT TCA CAC CCA GAG 1139
47   A   R   T   V   G   S   V   A   K   S   M   G   D   Y   L   D   S   H   P   E   66

1140 ACC AAC CAG GTG ATG ACC GCG GTC TTG CAG CAG CAG GTA GGG CCG GGG TCG GTC GCA TCG 1199
67   T   N   Q   V   M   T   A   V   L   Q   Q   Q   V   G   P   G   S   V   A   S   86

1200 CTG AAG GCC CAT TTC GAG GCG AAT CCC AAG GTC GCA TCG GAT CC 1243
87   L   K   A   H   F   E   A   N   P   K   V   A   S   D)   100

SEQ ID NOS. 1-7

FIG. 1

Insert of the clone containing DP428 and contained in seq1

```
1/1
  SEQ ID NO: 8                               31/11
GAT CGC CTT TGA CGC CTA TTC GGT CGC GCA     GCT TTT TGG CGA CGT CAC TGG TGC CCG CGT
asp arg leu OPA(arg leu phe gly arg ala     ala phe trp arg arg his trp cys pro arg
61/21           SEQ ID NO: 9                 91/31
TGC GGG CGT CCA GCC GCA GCG ACA CCA CAT     ACG GCC GGT CCG GTT GCG GGG GCC GTT GGG
cys gly arg pro ala ala ala thr pro his     thr ala gly pro val ala gly ala val gly
121/41                                       151/51
TGG GGT TGG GTG CCT CCG TCA CCC CAG GCA     GTT CGC TGG CTA TTT GTC GCA GTA GCG CGA
trp gly trp val pro pro ser pro gln ala     val arg trp leu phe val ala val ala arg
181/61                                       211/71
CGG CAT TGT CGA TGT CTT GGT AGC TAG CAT     CCG GTC GGG GGG CCG CTA CCA GCG CCA GCG
arg his cys arg cys leu gly ser)AMB(his     pro val gly gly pro leu pro ala pro ala
241/81           SEQ ID NO: 10               271/91
CCG GGG CTC CCC GGT CCG GGT AGT GCG CGT     CGA GTT GGT CGT GGA CCA GCA ATG ACT GCG
pro gly leu pro gly pro gly ser ala arg     arg val gly arg gly pro ala met thr ala
301/101                                      331/111
ACC CGG CGA CTT CGA AAC CGC CAC CGG TTA     GAT TCC CCG ACT GCG TCA TCG CCA GGT AAA
thr arg arg leu arg asn arg his arg leu     asp ser pro thr ala ser ser pro gly lys
361/121                                      391/131
CCG CCG GCA CTA ACG CCA GCA ACC AAC CCG     TGA AGA CCA ACC AAC GGC ACC TGC GCA GGT
pro pro ala leu thr pro ala thr asn pro)OPA(arg pro thr asn gly thr cys ala gly
421/141                                      451/151   SEQ ID NO: 11
TGC GGC TCA ACC GCA TCA TGA ACT GCT GGA     TTT CGG ACT CCC CGT ACT CTC GCG CAG TGC
cys gly ser thr ala ser)OPA(thr ala gly     phe arg thr pro arg thr leu ala gln cys
481/161          SEQ ID NO: 12               511/171
GTG CCC GCG AGC CTA CCG AAG ATC GCG TGC     ATG CGT TCG GCG TGG ACC GCA CAG CAC CTG
val pro ala ser leu pro lys ile ala cys     met arg ser ala trp thr ala gln his leu
541/181                                      571/191
GAG TTG GCG GCG CCG AGG GCC GAG ATG GCA     GGA TGA CGG ATC GTC GGG GGC GGG AAC TCC
glu leu ala ala pro arg ala glu met ala     gly)OPA(arg ile val gly gly gly asn ser
601/201                                      631/211   SEQ ID NO: 13
CAG GCC GCC GGA CCG TCG CAA ACC CGT CGC     AAA CCC GTC GCA AAC CGT AAG GAG TCA TCC
gln ala ala gly pro ser gln thr arg arg     lys pro val ala asn arg lys glu ser ser
661/221                                      691/231
ATG AAG ACA GGC ACC GCG ACG ACG CGG CGC     AGG CTG TTG GCA GTA CTG ATC GCC CTC GCG
met lys thr gly thr ala thr thr arg arg     arg leu leu ala val leu ile ala leu ala
721/241                                      751/251
TTG CCG GGG GCC GCC GTT GCG CTG CTG GCC     GAA CCA TCA GCG ACC GGC GCG TCG GAC CCG
leu pro gly ala ala val ala leu leu ala     glu pro ser ala thr gly ala ser asp pro
781/261                                      811/271
TGC GCG GCC AGC GAA GTG GCG AGG ACG GTC     GGT TCG GTC GCC AAG TCG ATG GGC GAC TAC
cys ala ala ser glu val ala arg thr val     gly ser val ala lys ser met gly asp tyr
841/281                                      871/291
CTG GAT TCA CAC CCA GAG ACC AAC CAG GTG     ATG ACC GCG GTC TTG CAG CAG CAG GTA GGG
leu asp ser his pro glu thr asn gln val     met thr ala val leu gln gln gln val gly
901/301                                      931/311
CCG GGG TCG GTC GCA TCG CTG AAG GCC CAT     TTC GAG GCG AAT CCC AAG GTC GCA TCG GAT C
pro gly ser val ala ser leu lys ala his     phe glu ala asn pro lys val ala ser asp)
```

SEQ ID NOS:8-13

FIG. 1A'

Insert of the clone containing DP428, other reading frame
2/1 ⟋SEQ ID NO: 14   32/11

ATC GCC TTT GAC GCC TAT TCG GTC GCG CAG CTT TTT GGC GAC GTC ACT GGT GCC CGC GTT
(ile ala phe asp ala tyr ser val ala gln leu phe gly asp val thr gly ala arg val
62/21 ⟍SEQ ID NO: 15   92/31

GCG GGC GTC CAG CCG CAG CGA CAC CAC ATA CGG CCG GTC CGG TTG CGG GGG CCG TTG GGT
ala gly val gln pro gln arg his his ile arg pro val arg leu arg gly pro leu gly
122/41   152/51

GGG GTT GGG TGC CTC CGT CAC CCC AGG CAG TTC GCT GGC TAT TTG TCG CAG TAG CGC GAC
gly val gly cys leu arg his pro arg gln phe ala gly tyr leu ser gln)AMB(arg asp
182/61   212/71   SEQ ID NO: 16⤵

GGC ATT GTC GAT GTC TTG GTA GCT AGC ATC CGG TCG GGG GGC CGC TAC CAG CGC CAG CGC
gly ile val asp val leu val ala ser ile arg ser gly gly arg tyr gln arg gln arg
242/81   272/91

CGG GGC TCC CCG GTC CGG GTA GTG CGC GTC GAG TTG GTC GTG GAC CAG CAA TGA CTG CGA
arg gly ser pro val arg val val arg val glu leu val val asp gln gln)OPA(leu arg
302/101   332/111   SEQ ID NO: 17⟋⤴

CCC GGC GAC TTC GAA ACC GCC ACC GGT TAG ATT CCC CGA CTG CGT CAT CGC CAG GTA AAC
pro gly asp phe glu thr ala thr gly)AMB(ile pro arg leu arg his arg gln val asn
362/121   SEQ ID NO: 18 ⟋⤴392/131

CGC CGG CAC TAA CGC CAG CAA CCA ACC CGT GAA GAC CAA CCA ACG GCA CCT GCG CAG GTT
arg arg his)OCH(arg gln gln pro thr arg glu asp gln pro thr ala pro ala gln val
422/141   ⟍SEQ ID NO: 19   452/151

GCG GCT CAA CCG CAT CAT GAA CTG CTG GAT TTC GGA CTC CCC GTA CTC TCG CGC AGT GCG
ala ala gln pro his his glu leu leu asp phe gly leu pro val leu ser arg ser ala
482/161   512/171

TGC CCG CGA GCC TAC CGA AGA TCG CGT GCA TGC GTT CGG CGT GGA CCG CAC AGC ACC TGG
cys pro arg ala tyr arg arg ser arg ala cys val arg arg gly pro his ser thr trp
542/181   572/191

AGT TGG CGG CGC CGA GGG CCG AGA TGG CAG GAT GAC GGA TCG TCG GGG GCG GGA ACT CCC
ser trp arg arg arg gly pro arg trp gln asp asp gly ser ser gly ala gly thr pro)
602/201   632/211

AGG CCG CCG GAC CGT CGC AAA CCC GTC GCA AAC CCG TCG CAA ACC GTA AGG AGT CAT CCA
arg pro pro asp arg arg lys pro val ala asn pro ser gln thr val arg ser his pro
662/221   692/231

TGA AGA CAG GCA CCG CGA CGA CGC GGC GCA GGC TGT TGG CAG TAC TGA TCG CCC TCG CGT
OPA(arg gln ala pro arg arg arg gly ala gly cys trp gln tyr)OPA(ser pro ser arg
722/241   ⟍SEQ ID NO: 20   752/251   SEQ ID NO: 21⟋⤴

TGC CGG GGG CCG CCG TTG CGC TGC TGG CCG AAC CAT CAG CGA CCG GCG CGT CGG ACC CGT
cys arg gly pro pro leu arg cys trp pro asn his gln arg pro ala arg arg thr arg
782/261   812/271

GCG CGG CCA GCG AAG TGG CGA GGA CGG TCG GTT CGG TCG CCA AGT CGA TGG GCG ACT ACC
ala arg pro ala lys trp arg gly arg ser val arg ser pro ser arg trp ala thr thr
842/281   872/291

TGG ATT CAC ACC CAG AGA CCA ACC AGG TGA TGA CCG CGG TCT TGC AGC AGC AGG TAG GGC
trp ile his thr gln arg pro thr arg)OPA OPA(pro arg ser cys ser ser arg)AMB(gly
902/301   SEQ ID NO: 22 ⟍932/311   SEQ ID NO: 23⟋⤴

CGG GGT CGG TCG CAT CGC TGA AGG CCC ATT TCG AGG CGA ATC CCA AGG TCG CAT CGG ATC
arg gly arg ser his arg)OPA(arg pro ile ser arg arg ile pro arg ser his arg ile)
                           ⟍SEQ ID NO: 24

SEQ ID NOS: 14-24

FIG. 1B'

Seq1C: Insert of the DP428 clone, other reading frame

3/1 **SEQ ID NO: 25**   33/11

TCG CCT TTG ACG CCT ATT CGG TCG CGC AGC   TTT TTG GCG ACG TCA CTG GTG CCC GCG TTG
(ser pro leu thr pro ile arg ser arg ser   phe leu ala thr ser leu val pro ala leu

63/21 **SEQ ID NO: 26**   93/31

CGG GCG TCC AGC CGC AGC GAC ACC ACA TAC   GGC CGG TCC GGT TGC GGG GGC CGT TGG GTG
arg ala ser ser arg ser asp thr thr tyr   gly arg ser gly cys gly gly arg trp val

123/41   153/51

GGG TTG GGT GCC TCC GTC ACC CCA GGC AGT   TCG CTG GCT ATT TGT CGC AGT AGC GCG ACG
gly leu gly ala ser val thr pro gly ser   ser leu ala ile cys arg ser ser ala thr

183/61   213/71

GCA TTG TCG ATG TCT TGG TAG CTA GCA TCC   GGT CGG GGG GCC GCT ACC AGC GCC AGC GCC
ala leu ser met ser trp)AMB(leu ala ser   gly arg gly ala ala thr ser ala ser ala

243/81 **SEQ ID NO: 27**   273/91

GGG GCT CCC CGG TCC GGG TAG TGC GCG TCG   AGT TGG TCG TGG ACC AGC AAT GAC TGC GAC
gly ala pro arg ser gly)AMB(cys ala ser   ser trp ser trp thr ser asn asp cys asp

303/101 **SEQ ID NO: 28**   333/111

CCG GCG ACT TCG AAA CCG CCA CCG GTT AGA   TTC CCC GAC TGC GTC ATC GCC AGG TAA ACC
pro ala thr ser lys pro pro pro val arg   phe pro asp cys val ile ala arg)OCH(thr

363/121   393/131 **SEQ ID NO: 29**

GCC GGC ACT AAC GCC AGC AAC CAA CCC GTG   AAG ACC AAC CAA CGG CAC CTG CGC AGG TTG
ala gly thr asn ala ser asn gln pro val   lys thr asn gln arg his leu arg arg leu

423/141   453/151

CGG CTC AAC CGC ATC ATG AAC TGC TGG ATT   TCG GAC TCC CCG TAC TCT CGC GCA GTG CGT
arg leu asn arg ile met asn cys trp ile   ser asp ser pro tyr ser arg ala val arg

483/161   513/171

GCC CGC GAG CCT ACC GAA GAT CGC GTG CAT   GCG TTC GGC GTG GAC CGC ACA GCA CCT GGA
ala arg glu pro thr glu asp arg val his   ala phe gly val asp arg thr ala pro gly

543/181   573/191

GTT GGC GGC GCC GAG GGC CGA GAT GGC AGG   ATG ACG GAT CGT CGG GGG CGG GAA CTC CCA
val gly gly ala glu gly arg asp gly arg   met thr asp arg arg gly arg glu leu pro

603/201   633/211

GGC CGC CGG ACC GTC GCA AAC CCG TCG CAA   ACC CGT CGC AAA CCG TAA GGA GTC ATC CAT
gly arg arg thr val ala asn pro ser gln   thr arg arg lys pro)OCH(gly val ile his

663/221   693/231 **SEQ ID NO: 30**

GAA GAC AGG CAC CGC GAC GAC GCG GCG CAG   GCT GTT GGC AGT ACT GAT CGC CCT CGC GTT
glu asp arg his arg asp asp ala ala gln   ala val gly ser thr asp arg pro arg val

723/241   753/251

GCC GGG GGC CGC CGT TGC GCT GCT GGC CGA   ACC ATC AGC GAC CGG CGC GTC GGA CCC GTG
ala gly gly arg arg cys ala ala gly arg   thr ile ser asp arg arg val gly pro val

783/261   813/271

CGC GGC CAG CGA AGT GGC GAG GAC GGT CGG   TTC GGT CGC CAA GTC GAT GGG CGA CTA CCT
arg gly gln arg ser gly glu asp gly arg   phe gly arg gln val asp gly arg leu pro

843/281   873/291

GGA TTC ACA CCC AGA GAC CAA CCA GGT GAT   GAC CGC GGT CTT GCA GCA GCA GGT AGG GCC
gly phe thr pro arg asp gln pro gly asp   asp arg gly leu ala ala ala gly arg ala

903/301   933/311

GGG GTC GGT CGC ATC GCT GAA GGC CCA TTT   CGA GGC GAA TCC CAA GGT CGC ATC GGA TC
gly val gly arg ile ala glu gly pro phe   arg gly glu ser gln gly arg ile gly)

SEQ ID NOS: 25-30

FIG. 1C'

Coding sequence DP428 identical to the Rv0203 predicted by Cole et al.
(Nature 393:537-544)

1/1 ⟋ SEQ ID NO: 31   31/11
ᵞATG AAG ACA GGC ACC GCG ACG ACG CGG CGC AGG CTG TTG GCA GTA CTG ATC GCC CTC GCG
(Met lys thr gly thr ala thr thr arg arg arg leu leu ala val leu ile ala leu ala

61/21 ⟍ SEQ ID NO: 32   91/31
TTG CCG GGG GCC GCC GTT GCG CTG CTG GCC GAA CCA TCA GCG ACC GGC GCG TCG GAC CCG
leu pro gly ala ala val ala leu leu ala glu pro ser ala thr gly ala ser asp pro

121/41   151/51
TGC GCG GCC AGC GAA GTG GCG AGG ACG GTC GGT TCG GTC GCC AAG TCG ATG GGC GAC TAC
cys ala ala ser glu val ala arg thr val gly ser val ala lys ser met gly asp tyr

181/61   211/71
CTG GAT TCA CAC CCA GAG ACC AAC CAG GTG ATG ACC GCG GTC TTG CAG CAG CAG GTA GGG
leu asp ser his pro glu thr asn gln val met thr ala val leu gln gln gln val gly

241/81   271/91
CCG GGG TCG GTC GCA TCG CTG AAG GCC CAT TTC GAG GCG AAT CCC AAG GTC GCA TCG GAT
pro gly ser val ala ser leu lys ala his phe glu ala asn pro lys val ala ser asp

301/101   331/111
CTG CAC GCG CTT TCG CAA CCG CTG ACC GAT CTT TCG ACT CGG TGC TCG CTG CCG ATC AGC
leu his ala leu ser gln pro leu thr asp leu ser thr arg cys ser leu pro ile ser

361/121   391/131
GGC CTG CAG GCG ATC GGT TTG ATG CAG GCG GTG CAG GGC GCC CGC CGG TAG
gly leu gln ala ile gly leu met gln ala val gln gly ala arg arg) AMB

SEQ ID NOS:31-32

FIG. 1D'

ORF containing the DP428 sequence and forming part of seqlA'

1/1 ⟋ SEQ ID NO: 33   31/11
TGA CGG ATC GTC GGG GGC GGG AAC TCC CAG GCC GCC GGA CCG TCG CAA ACC CGT CGC AAA
OPA(arg ile val gly gly gly asn ser gln ala ala gly pro ser gln thr arg arg lys

61/21 ⟍ SEQ ID NO: 34   91/31
CCC GTC GCA AAC CGT AAG GAG TCA TCC ATG AAG ACA GGC ACC GCG ACG ACG CGG CGC AGG
pro val ala asn arg lys glu ser ser met lys thr gly thr ala thr thr arg arg arg

121/41   151/51
CTG TTG GCA GTA CTG ATC GCC CTC GCG TTG CCG GGG GCC GCC GTT GCG CTG CTG GCC GAA
leu leu ala val leu ile ala leu ala leu pro gly ala ala val ala leu leu ala glu

181/61   211/71
CCA TCA GCG ACC GGC GCG TCG GAC CCG TGC GCG GCC AGC GAA GTG GCG AGG ACG GTC GGT
pro ser ala thr gly ala ser asp pro cys ala ala ser glu val ala arg thr val gly

241/81   271/91
TCG GTC GCC AAG TCG ATG GGC GAC TAC CTG GAT TCA CAC CCA GAG ACC AAC CAG GTG ATG
ser val ala lys ser met gly asp tyr leu asp ser his pro glu thr asn gln val met

301/101   331/111
ACC GCG GTC TTG CAG CAG CAG GTA GGG CCG GGG TCG GTC GCA TCG CTG AAG GCC CAT TTC
thr ala val leu gln gln gln val gly pro gly ser val ala ser leu lys ala his phe

361/121   391/131
GAG GCG AAT CCC AAG GTC GCA TCG GAT CTG CAC GCG CTT TCG CAA CCG CTG ACC GAT CTT
glu ala asn pro lys val ala ser asp leu his ala leu ser gln pro leu thr asp leu

421/141   451/151
TCG ACT CGG TGC TCG CTG CCG ATC AGC GGC CTG CAG GCG ATC GGT TTG ATG CAG GCG GTG
ser thr arg cys ser leu pro ile ser gly leu gln ala ile gly leu met gln ala val

481/161
CAG GGC GCC CGC CGG TAG
gln gly ala arg arg)AMB

SEQ ID NOS:33-34

FIG. 1E'

```
491 CCGGTCGGGGGGCCGCTACCAGCGCCAGCGCCGGGGCTCCCCGGTCCGGGTA GTG CGC GTC GAG TTG GTC GTG 563
  1 SEQ ID NO:  35              SEQ ID NO:  37 ──────────► (V   R   V   E   L   V   V   7

564 GAC CAG CAA TGA CTGCGACCCGGCGACTTCGAAACCGCCACCGGTTAGATTCCCCGACTGCGTCATCGCCAGGTAA 639
  8 D   Q   Q)   *  SEQ ID NO:  36

640 ACCGCCGGCACTAACGCCAGCAACCAACCC GTG AAG ACC AAC CAA CGG CAC CTG CGC AGG TTG CGG 705
  1  SEQ ID NO:  38 ──────────► (V   K   T   N   Q   R   H   L   R   R   L   R  12

706 CTC AAC CGC ATC ATG AAC TGC TGG ATT TCG GAC TCC CCG TAC TCT CGC GCA GTG CGT GCC 765
 13 L   N   R   I   M   N   C   W   I   S   D   S   P   Y   S   R   A   V   R   A  32

766 CGC GAG CCT ACC GAA GAT CGC GTG CAT GCG TTC GGC GTG GAC CGC ACA GCA CCT GGA GTT 825
 33 R   E   P   T   E   D   R   V   H   A   F   G   V   D   R   T   A   P   G   V  52

826 GGC GGC GCC GAG GGC CGA GAT GGC AGG ATG ACG GAT CGT CGG GGG CGG GAA CTC CCA GGC 885
 53 G   G   A   E   G   R   D   G   R   M   T   D   R   R   G   R   E   L   P   G  72

886 CGC CGG ACC GTC GCA AAC CCG TCG CAA ACC CGT CGC AAA CCG TAA GGAGTCATCC ATG AAG 946
 73 R   R   T   V   A   N   P   S   Q   T   R   R   K   P)  *  xxxxxx    ──►(M   K   2
                                                              SEQ ID NO:  39

947 ACA GGC ACC GCG ACG ACG CGG CGC AGG CTG TTG GCA GTA CTG ATC GCC CTC GCG TTG CCG 1006
  3 T   G   T   A   T   T   R   R   R   L   L   A   V   L   I   A   L   A   L   P  22

1007 GGG GCC GCC GTT GCG CTG CTG GCC GAA CCA TCA GCG ACC GGC GCG TCG GAC CCG TGC GCG 1066
  23 G   A   A   V   A   L   L   A   E   P   S   A   T   G   A   S   D   P   C   A  42

1067 GCC AGC GAA GTG GCG AGG ACG GTC GGT TCG GTC GCC AAG TCG ATG GGC GAC TAC CTG GAT 1126
  43 A   S   E   V   A   R   T   V   G   S   V   A   K   S   M   G   D   Y   L   D  62

1127 TCA CAC CCA GAG ACC AAC CAG GTG ATG ACC GCG GTC TTG CAG CAG CAG GTA GGG CCG GGG 1186
  63 S   H   P   E   T   N   Q   V   M   T   A   V   L   Q   Q   Q   V   G   P   G  82

1187 TCG GTC GCA TCG CTG AAG GCC CAT TTC GAG GCG AAT CCC AAG GTC GCA TCG GAT CTG CAC 1246
  83 S   V   A   S   L   K   A   H   F   E   A   N   P   K   V   A   S   D   L   H  102

1247 GCG CTT TCG CAA CCG CTG ACC GAT CTT TCG ACT CGG TGC TCG CTG CCG ATC AGC GGC CTG 1306
 103 A   L   S   Q   P   L   T   D   L   S   T   R   C   S   L   P   I   S   G   L  122

1307 CAG GCG ATC GGT TTG ATG CAG GCG GTG CAG GGC GCC CGC CGG TAG  ATG CCG GAC CGC CGC 1366
 123 Q   A   I   G   L   M   Q   A   V   Q   G   A   R   R)  *   (M   P   D   R   R   5
                                                                 SEQ ID NO:  40

1367 CGG GTC CGG CGC AGT CGA CGT GAG GCA GCG GTC GCC TAC CGG GGC GGT GTC TCG CCG CCT 1426
   6 R   V   R   R   S   R   R   E   A   A   V   A   Y   R   G   G   V   S   P   P  25

1427 TCT GGT CGC AGG TCA GGG GTC GGC GCT GGA CCT TGC GGT GTG GTT TCG ACC GGG TCG TCG 1486
  26 S   G   R   R   S   G   V   G   A   G   P   C   G   V   V   S   T   G   S   S  45

1487 CAG GGT GTG CCC TGC GGT TGG ATG ACA AGT CGC AGG TTT GGA TCG GTT GGC GGG TCG CGA 1546
  46 Q   G   V   P   C   G   W   M   T   S   R   R   F   G   S   V   G   G   S   R  65

1547 TCG TTG T                                                                     1553
  66 S   L)                                                                        67
```

SEQ ID NOS:35-40

FIG. 2

```
SEQ ID NO:   41                                    31/11
TCG CCG GCT CGC GGA CGT AGA TAA TAG CTC ACC GTT GGA CGA CCT CGA CAG GGT CCT TTG
(ser pro ala arg gly arg arg)OCH AMB(leu thr val gly arg pro arg gln gly pro leu)
61/21    SEQ ID NO:   42    SEQ ID NO: 43    91/31
TGA CTG CCG GGC TTG ACG CGG ACG ACC ACA GAG TCG GGT CAT CGC CTA AGG CTA CCG TTC
OPA(leu pro gly leu thr arg thr thr thr glu ser gly his arg leu arg leu pro phe)
121/41    SEQ ID NO:   44                        151/51
TGA CCT GGG GTG CGT GGG CGC CGA CGA GTG AGG CAG TCA TGT CTC AGG GCC CAC CGC CAC
OPA(pro gly val arg gly arg arg arg val arg gln ser cys leu arg ala his arg his
181/61    SEQ ID NO:   45                        211/71
CTC GGT CGC CGG CAG TGT CAG CAT GTG CAG ATG ACT CCA CGC AGC TTG TTC GTG TTG GTG
leu gly arg arg gln cys gln his val gln met thr pro arg ser leu phe val leu val
241/81                                          271/91
TCG TGG TTG CGA CGA CTT GGC GCT GGT GAG CGC ACC CGC CGG CGT CGT GCC GCG CAT GCG
ser trp leu arg arg leu gly ala gly glu arg thr arg arg arg arg ala ala his ala
301/101
GAT C
asp)
```

SEQ ID NOS:41-45

FIG. 3A

```
SEQ ID NO:   46                                    32/11
CGC CGG CTC GCG GAC GTA GAT AAT AGC TCA CCG TTG GAC GAC CTC GAC AGG GTC CTT TGT
(arg arg leu ala asp val asp asn ser ser pro leu asp asp leu asp arg val leu cys
62/21    SEQ ID NO:   47                        92/31
GAC TGC CGG GCT TGA CGC GGA CGA CCA CAG AGT CGG GTC ATC GCC TAA GGC TAC CGT TCT
asp cys arg ala)OPA(arg gly arg pro gln ser arg val ile ala)OCH(gly tyr arg ser
122/41              SEQ ID NO:   48    152/51              SEQ ID NO:   49
GAC CTG GGG TGC GTG GGC GCC GAC GAG TGA GGC AGT CAT GTC TCA GGG CCC ACC GCC ACC
asp leu gly cys val gly ala asp glu)OPA(gly ser his val ser gly pro thr ala thr
182/61    SEQ ID NO:   50    212/71
TCG GTC GCC GGC AGT GTC AGC ATG TGC AGA TGA CTC CAC GCA GCT TGT TCG TGT TGG TGT
ser val ala gly ser val ser met cys arg)OPA(leu his ala ala cys ser cys trp cys
242/81                                272/91    SEQ ID NO:   51
CGT GGT TGC GAC GAC TTG GCG CTG GTG AGC GCA CCC GCC GGC GTC GTG CCG CGC ATG CGG
arg gly cys asp asp leu ala leu val ser ala pro ala gly val val pro arg met arg
302/101
ATC
ile)
```

SEQ ID NOS:46-51

FIG. 3B

SEQ ID NO: 52   33/11
GCC GGC TCG CGG ACG TAG ATA ATA GCT CAC CGT TGG ACG ACC TCG ACA GGG TCC TTT GTG
(ala gly ser arg thr)AMB(ile ile ala his arg trp thr thr ser thr gly ser phe val
63/21   SEQ ID NO: 53   SEQ ID NO: 54 93/31
ACT GCC GGG CTT GAC GCG GAC GAC CAC AGA GTC GGG TCA TCG CCT AAG GCT ACC GTT CTG
thr ala gly leu asp ala asp asp his arg val gly ser ser pro lys ala thr val leu
123/41   153/51
ACC TGG GGT GCG TGG GCG CCG ACG AGT GAG GCA GTC ATG TCT CAG GGC CCA CCG CCA CCT
thr trp gly ala trp ala pro thr ser glu ala val met ser gln gly pro pro pro pro
183/61   213/71
CGG TCG CCG GCA GTG TCA GCA TGT GCA GAT GAC TCC ACG CAG CTT GTT CGT GTT GGT GTC
arg ser pro ala val ser ala cys ala asp asp ser thr gln leu val arg val gly val
243/81   273/91
GTG GTT GCG ACG ACT TGG CGC TGG TGA GCG CAC CCG CCG GCG TCG TGC CGC GCA TGC GGA
val val ala thr thr trp arg trp)OPA(ala his pro pro ala ser cys arg ala cys gly)
                                         SEQ ID NO: 55
TC

SEQ ID NOS:52-55

FIG. 3C

SEQ ID NO: 56   31/11
CCA ATT TTC CTT CGC GCC GTG CAA TAC CAT CTG CAA GAC CAG CGA CGG CCC GTG GTT GCG
(pro ile phe leu arg ala val gln tyr his leu gln asp gln arg arg pro val val ala
61/21   SEQ ID NO: 57   91/31
GTC GCG CAG CTT GCG GAA ACC GGG TAT GGA CCC TGC CGT ACC GTT GTT GCC ACT TGA TGT
val ala gln leu ala glu thr gly tyr gly pro cys arg thr val val ala thr)OPA(cys
121/41   151/51   SEQ ID NO: 58
CGT CGC TCT CCA CCC GTC GGG GGG CGA AAG CCA TTC CGA CAC TGG GAT CCT CAA AAC GTC
arg arg ser pro pro val gly gly arg lys pro phe arg his trp asp pro gln asn val
181/61   211/71
GGC TGA GTG TCT GCA GGG CTC CGG GGA GCA GCC GAT CAT CAC CAT GTA CGA ACT GAA TAA
gly)OPA(val ser ala gly leu arg gly ala ala asp his his his val arg thr glu)OCH
241/81   SEQ ID NO: 59   271/91
GTC CCC CGC GCG CGA CTT CCA GAC ATT TGT TGT GGT TTC GGT TGA GGC CGA GGC GAG GCT
(val pro arg ala arg leu pro asp ile cys cys gly phe gly)OPA(gly arg gly glu ala
301/101   SEQ ID NO: 60   331/111   SEQ ID NO: 61
CAT TTC GCA GCA ACC GGT CTC CGG GTC GCA GCA TCG TTG CGG CGA TCG CGG CGC AGT CGT
his phe ala ala thr gly leu arg val ala ala ser leu arg arg ser arg arg ser arg
361/121
CGG ACG AGT CGT CGT CAA CGA CCA CGA TC
arg thr ser arg arg gln arg pro arg)

SEQ ID NOS:56-61

FIG. 4A

SEQ ID NO: 62   32/11
CAA TTT TCC TTC GCG CCG TGC AAT ACC ATC TGC AAG ACC AGC GAC GGC CCG TGG TTG CGG
gln(phe ser phe ala pro cys asn thr ile cys lys thr ser asp gly pro trp leu arg
62/21 SEQ ID NO: 63   92/31
TCG CGC AGC TTG CGG AAA CCG GGT ATG GAC CCT GCC GTA CCG TTG TTG CCA CTT GAT GTC
ser arg ser leu arg lys pro gly met asp pro ala val pro leu leu pro leu asp val
122/41   152/51
GTC GCT CTC CAC CCG TCG GGG GGC GAA AGC CAT TCC GAC ACT GGG ATC CTC AAA ACG TCG
val ala leu his pro ser gly gly glu ser his ser asp thr gly ile leu lys thr ser
182/61   212/71
GCT GAG TGT CTG CAG GGC TCC GGG GAG CAG CCG ATC ATC ACC ATG TAC GAA CTG AAT AAG
ala glu cys leu gln gly ser gly glu gln pro ile ile thr met tyr glu leu asn lys
242/81   272/91
TCC CCC GCG CGC GAC TTC CAG ACA TTT GTT GTG GTT TCG GTT GAG GCC GAG GCG AGG CTC
ser pro ala arg asp phe gln thr phe val val val ser val glu ala glu ala arg leu
302/101   332/111
ATT TCG CAG CAA CCG GTC TCC GGG TCG CAG CAT CGT TGC GGC GAT CGC GGC GCA GTC GTC
ile ser gln gln pro val ser gly ser gln his arg cys gly asp arg gly ala val val
362/121
GGA CGA GTC GTC GTC AAC GAC CAC GAT C
gly arg val val val asn asp his asp)

SEQ ID NOS:62-63

FIG. 4B

SEQ ID NO:64   33/11
AAT TTT CCT TCG CGC CGT GCA ATA CCA TCT GCA AGA CCA GCG ACG GCC CGT GGT TGC GGT
(asn phe pro ser arg arg ala ile pro ser ala arg pro ala thr ala arg gly cys gly
63/21 SEQ ID NO: 65   93/31
CGC GCA GCT TGC GGA AAC CGG GTA TGG ACC CTG CCG TAC CGT TGT TGC CAC TTG ATG TCG
arg ala ala cys gly asn arg val trp thr leu pro tyr arg cys cys his leu met ser
123/41   153/51
TCG CTC TCC ACC CGT CGG GGG GCG AAA GCC ATT CCG ACA CTG GGA TCC TCA AAA CGT CGG
ser leu ser thr arg arg gly ala lys ala ile pro thr leu gly ser ser lys arg arg
183/61   213/71
CTG AGT GTC TGC AGG GCT CCG GGG AGC AGC CGA TCA TCA CCA TGT ACG AAC TGA ATA AGT
leu ser val cys arg ala pro gly ser ser arg ser ser pro cys thr asn)OPA(ile ser
243/81   273/91 SEQ ID NO: 66
CCC CCG CGC GCG ACT TCC AGA CAT TTG TTG TGG TTT CGG TTG AGG CCG AGG CGA GGC TCA
pro pro arg ala thr ser arg his leu leu trp phe arg leu arg pro arg arg gly ser
303/101   333/111
TTT CGC AGC AAC CGG TCT CCG GGT CGC AGC ATC GTT GCG GCG ATC GCG GCG CAG TCG TCG
phe arg ser asn arg ser pro gly arg ser ile val ala ala ile ala ala gln ser ser
363/121
GAC GAG TCG TCG TCA ACG ACC ACG ATC
asp glu ser ser ser thr thr thr ile)

SEQ ID NOS:64-66

FIG.4C

part of the nucleotide sequence of seq4A

1/1 SEQ ID NO: 67   31/11
CCG CGC GCG ACT TCC AGA CAT TTG TTG TGG TTT CGG TTG AGG CCG AGG CGA GGC TCA TTT
(pro arg ala thr ser arg his leu leu trp phe arg leu arg pro arg arg gly ser phe
61/21   SEQ ID NO: 68   91/31
CGC AGC AAG CGG TCT CCG GGT CGC AGC ATC GTT GCG GCG ATC GCG GCG CAG TCG TCG GAC
arg ser lys arg ser pro gly arg ser ile val ala ala ile ala ala gln ser ser asp
121/41
GAG TCG TCG TCA ACG ACC ACG ATC
glu ser ser ser thr thr thr ile)

SEQ ID NOS:67-68

FIG.4A'

1/1 SEQ ID NO: 69   31/11
CGC GCG CGA CTT CCA GAC ATT TGT TGT GGT TTC GGT TGA GGC CGA GGC GAG GCT CAT TTC
(arg ala arg leu pro asp ile cys cys gly phe gly)OPA(gly arg gly glu ala his phe
61/21   SEQ ID NO: 70   91/31   SEQ ID NO: 71
GCA GCA AGC GGT CTC CGG GTC GCA GCA TCG TTG CGG CGA TCG CGG CGC AGT CGT CGG ACG
ala ala ser gly leu arg val ala ala ser leu arg arg ser arg arg ser arg arg thr
121/41
AGT CGT CGT CAA CGA CCA CGA TC
ser arg arg gln arg pro arg)

SEQ ID NOS:69-71

FIG.4B'

1/1 SEQ ID NO: 72   31/11
GCC GCG CGC GAC TTC CAG ACA TTT GTT GTG GTT TCG GTT GAG GCC GAG GCG AGG CTC ATT
(ala ala arg asp phe gln thr phe val val val ser val glu ala glu ala arg leu ile
61/21 SEQ ID NO: 73   91/31
TCG CAG CAA GCG GTC TCC GGG TCG CAG CAT CGT TGC GGC GAT CGC GGC GCA GTC GTC GGA
ser gln gln ala val ser gly ser gln his arg cys gly asp arg gly ala val val gly
121/41
CGA GTC GTC GTC AAC GAC CAC GAT C
arg val val val asn asp his asp

SEQ ID NOS:72-73

FIG. 4C'

ORF according to Cole et al. (Nature 393:537-544) and containing the
sequence Seq 4A'
1/1 ___ SEQ ID NO: 74   31/11
tga ata agt ccg ccg cgc gcg act tcc aga cat ttg ttg tgg ttt cgg ttg agg ccg agg
OPA(ile ser pro pro arg ala thr ser arg his leu leu trp phe arg leu arg pro arg
61/21 SEQ ID NO: 75   91/31
cga ggc tca ttt cgc agc aag cgg tct ccg ggt cgc agc atc gtt gcg gcg atc gcg gcg
arg gly ser phe arg ser lys arg ser pro gly arg ser ile val ala ala ile ala ala
121/41   151/51
cag tcg tcg gac gag tcg tcg tca acg acc acg atc tcg aac tcg acg ccc tcc tgt tcg
gln ser ser asp glu ser ser ser thr thr thr ile ser asn ser thr pro ser cys ser
181/61   211/71
agg atg cta cgc aga cag cgc tcg atg gtg gcg ccg ttg ttg tac atc ggg atg cac acc
arg met leu arg arg gln arg ser met val ala pro leu leu tyr ile gly met his thr
241/81   271/91
gag ata agc ggt ttc gcc ggg ttc acc gat acc acg ctt gat gca tca cca ggc acc aca
glu ile ser gly phe ala gly phe thr asp thr thr leu asp ala ser pro gly thr thr
301/101
tgg cga ctc aga gac tag
trp arg leu arg asp) AMB

SEQ ID NOS:74-75

FIG. 4D'

sequence upstream of seq4A' and fused with seq4A'
1/1 ___ SEQ ID NO: 76   31/11
GCA ACC TAC CAG CAG AGC CAG GGG CTC ACA GGA CCT AAA GGA GTA GCG CCC ATG GCT GAT
(ala thr tyr gln gln ser gln gly leu thr gly pro lys gly val ala pro met ala asp)
    ____ SEQ ID NO: 77
C

SEQ ID NOS:76-77

FIG. 4E'

seq4J' in another reading frame
1/1 ___ SEQ ID NO: 78   31/11
ACG CAA CCT ACC AGC AGA GCC AGG GGC TCA CAG GAC CTA AAG GAG TAG CGC CCA TGG CTG
(thr gln pro thr ser arg ala arg gly ser gln asp leu lys glu)AMB(arg pro trp leu
61/21 ___ SEQ ID NO: 79   SEQ ID NO: 80 ___
ATC
ile)

SEQ ID NOS:78-80

FIG. 4F'

seq 4J' in the third reading frame
1/1 ___ SEQ ID NO: 81   31/11
CGC AAC CTA CCA GCA GAG CCA GGG GCT CAC AGG ACC TAA AGG AGT AGC GCC CAT GGC TGA
(arg asn leu pro ala glu pro gly ala his arg thr)OCH(arg ser ser ala his gly)OPA
    ___ SEQ ID NO: 82   SEQ ID NO: 83 ___
TC

SEQ ID NOS:81-83
FIG. 4G'

sequence Rv2050 predicted by Cole et al. (Nature 393:537-544) and
containing seq4J

```
1/1  SEQ ID NO:  84                      31/11
ATG GCT GAT CGT GTC CTG AGG GGC AGT CGC CTC GGA GCC GTG AGC TAT GAG ACC GAC CGC
(Met ala asp arg val leu arg gly ser arg leu gly ala val ser tyr glu thr asp arg
61/21       SEQ ID NO:  85                      91/31
AAC CAC GAC CTG GCG CCG CGC CAG ATC GCG CGG TAC CGC ACC GAC AAC GGC GAG GAG TTC
asn his asp leu ala pro arg gln ile ala arg tyr arg thr asp asn gly glu glu phe
121/41                                   151/51
GAA GTC CCG TTC GCC GAT GAC GCC GAG ATC CCC GGC ACC TGG TTG TGC CGC AAC GGC ATG
glu val pro phe ala asp asp ala glu ile pro gly thr trp leu cys arg asn gly met
181/61                                   211/71
GAA GGC ACC CTG ATC GAG GGC GAC CTG CCC GAG CCG AAG AAG GTT AAG CCG CCC CGG ACG
glu gly thr leu ile glu gly asp leu pro glu pro lys lys val lys pro pro arg thr
241/81                                   271/91
CAC TGG GAC ATG CTG CTG GAG CGC CGT TCC ATC GAA GAA CTC GAA GAG TTA CTT AAG GAG
his trp asp met leu leu glu arg arg ser ile glu glu leu glu glu leu leu lys glu
301/101                                  331/111
CGC CTC GAG CTC ATT CGG TCA CGT CGG CGC GGC TGA
arg leu glu leu ile arg ser arg arg arg gly)OPA
```

SEQ ID NOS:84-85

FIG. 4H'

ORF according to Cole et al. (Nature 393:537-544) and containing the
sequence Rv2050

```
1/1  SEQ ID NO:  86                      31/11
TAG TCC GCC CGG GTG TCC GAT CCC GGT ATC ATT GAT GGT CGC GCC GCG CGC GTC GCG TGC
AMB(ser ala arg val ser asp pro gly ile ile asp gly arg ala ala arg val ala cys
61/21       SEQ ID NO:  87                      91/31
CGG GAA CTA CGC AGA CGG CCG CAG CGT TTG CCA ACC GGA GCC AGT CGC CAG TAC GCA ACC
arg glu leu arg arg arg pro gln arg leu pro thr gly ala ser arg gln tyr ala thr
121/41                                   151/51
TAC CAG CAG AGC CCA GGG CTC ACA GGA CCT AAA GGA GTA GCG CCC ATG GCT GAT CGT GTC
tyr gln gln ser pro gly leu thr gly pro lys gly val ala pro met ala asp arg val
181/61                                   211/71
CTG AGG GGC AGT CGC CTC GGA GCC GTG AGC TAT GAG ACC GAC CGC AAC CAC GAC CTG GCG
leu arg gly ser arg leu gly ala val ser tyr glu thr asp arg asn his asp leu ala
241/81                                   271/91
CCG CGC CAG ATC GCG CGG TAC CGC ACC GAC AAC GGC GAG GAG TTC GAA GTC CCG TTC GCC
pro arg gln ile ala arg tyr arg thr asp asn gly glu glu phe glu val pro phe ala
301/101                                  331/111
GAT GAC GCC GAG ATC CCC GGC ACC TGG TTG TGC CGC AAC GGC ATG GAA GGC ACC CTG ATC
asp asp ala glu ile pro gly thr trp leu cys arg asn gly met glu gly thr leu ile
361/121                                  391/131
GAG GGC GAC CTG CCC GAG CCG AAG AAG GTT AAG CCG CCC CGG ACG CAC TGG GAC ATG CTG
glu gly asp leu pro glu pro lys lys val lys pro pro arg thr his trp asp met leu
421/141                                  451/151
CTG GAG CGC CGT TCC ATC GAA GAA CTC GAA GAG TTA CTT AAG GAG CGC CTC GAG CTC ATT
leu glu arg arg ser ile glu glu leu glu glu leu leu lys glu arg leu glu leu ile
481/161
CGG TCA CGT CGG CGC GGC TGA
arg ser arg arg arg gly)OPA
```

SEQ ID NOS 86-87

FIG. 4I'

SEQ ID NO: 88   31/11
GAT CGC GGT CAA CGA GGC CGA ATA CGG CGA GAT GTG GGC CCA AGA CGC CGC CGC GAT GTT
(asp arg gly gln arg gly arg ile arg arg asp val gly pro arg arg arg arg asp val
61/21   SEQ ID NO: 89   91/31
TGG CTA CGC CGC GGC GAC GGC GAC GGC GAC GGC GAC GTT GCT GCC GTT CGA GGA GGC GCC
trp leu arg arg gly asp gly asp gly asp gly asp val ala ala val arg gly gly ala
121/41   151/51
GGA GAT GAC CAG CGC GGG TGG GCT CCT CGA GCA GGC CGC CGC GGT CGA GGA GGC CTC CGA
gly asp asp gln arg gly trp ala pro arg ala gly arg arg gly arg gly gly leu arg
181/61   211/71
CAC CGC CGC GGC GAA CCA GTT GAT GAA CAA TGT GCC CCA GGC GCT GCA ACA GCT GGC CCA
his arg arg gly glu pro val asp glu gln cys ala pro gly ala ala thr ala gly pro
241/81   271/91
GCC CAC GCA GGG CAC CAC GCC TTC TTC CAA GCT GGG TGG CCT GTG GAA GAC GGT CTC GCC
ala his ala gly his his ala phe phe gln ala gly trp pro val glu asp gly leu ala
301/101
GCA TCG GTC GCC GAT C
ala ser val ala asp)

SEQ ID NOS.88-89

FIG. 5A

SEQ ID NO: 90   32/11
ATC GCG GTC AAC GAG GCC GAA TAC GGC GAG ATG TGG GCC CAA GAC GCC GCC GCG ATG TTT
ile ala val asn glu ala glu tyr gly glu met trp ala gln asp ala ala ala met phe
62/21   SEQ ID NO: 91   92/31
GGC TAC GCC GCG GCG ACG GCG ACG GCG ACG GCG ACG TTG CTG CCG TTC GAG GAG GCG CCG
gly tyr ala ala ala thr ala thr ala thr ala thr leu leu pro phe glu glu ala pro
122/41   152/51
GAG ATG ACC AGC GCG GGT GGG CTC CTC GAG CAG GCC GCC GCG GTC GAG GAG GCC TCC GAC
glu met thr ser ala gly gly leu leu glu gln ala ala ala val glu glu ala ser asp
182/61   212/71
ACC GCC GCG GCG AAC CAG TTG ATG AAC AAT GTG CCC CAG GCG CTG CAA CAG CTG GCC CAG
thr ala ala ala asn gln leu met asn asn val pro gln ala leu gln gln leu ala gln
242/81   272/91
CCC ACG CAG GGC ACC ACG CCT TCT TCC AAG CTG GGT GGC CTG TGG AAG ACG GTC TCG CCG
pro thr gln gly thr thr pro ser ser lys leu gly gly leu trp lys thr val ser pro
302/101
CAT CGG TCG CCG ATC
his arg ser pro ile

SEQ ID NOS.90-91

FIG. 5B

SEQ ID NO: 92   33/11
TCG CGG TCA ACG AGG CCG AAT ACG GCG AGA TGT GGG CCC AAG ACG CCG CCG CGA TGT TTG
ser(arg ser thr arg pro asn thr ala arg cys gly pro lys thr pro pro arg cys leu
63/21 SEQ ID NO: 93   93/31
GCT ACG CCG CGG CGA CGG CGA CGG CGA CGG CGA CGT TGC TGC CGT TCG AGG AGG CGC CGG
ala thr pro arg arg arg arg arg arg arg arg arg cys cys arg ser arg arg arg arg
123/41   153/51
AGA TGA CCA GCG CGG GTG GGC TCC TCG AGC AGG CCG CCG CGG TCG AGG AGG CCT CCG ACA
arg)OPA(pro ala arg val gly ser ser ser arg pro pro arg ser arg arg pro pro thr
183/61   SEQ ID NO: 94   213/71
CCG CCG CGG CGA ACC AGT TGA TGA ACA ATG TGC CCC AGG CGC TGC AAC AGC TGG CCC AGC
pro pro arg arg thr ser)OPA OPA(thr met cys pro arg arg cys asn ser trp pro ser
243/81   SEQ ID NO: 95   273/91
CCA CGC AGG GCA CCA CGC CTT CTT CCA AGC TGG GTG GCC TGT GGA AGA CGG TCT CGC CGC
pro arg arg ala pro arg leu leu pro ser trp val ala cys gly arg arg ser arg arg
303/101
ATC GGT CGC CGA TC
ile gly arg arg)

SEQ ID NOS:92-95

FIG. 5C

part of the nucleotide sequence Seq 5A

1/1 SEQ ID NO: 96   31/11
CGC CGC GGC GAC GGC GAC GGC GAC GGC GAC GTT GCT GCC GTT CGA GGA GGC GCC GGA GAT
arg arg gly asp gly asp gly asp gly asp val ala ala val arg gly gly ala gly asp
61/21 SEQ ID NO: 97   91/31
GAC CAG CGC GGG TGG GCT CCT CGA GCA GGC CGC CGC GGT CGA GGA GGC CTC CGA CAC CGC
asp gln arg gly trp ala pro arg ala gly arg arg gly arg gly gly leu arg his arg
121/41   151/51
CGC GGC GAA CCA GTT GAT GAA CAA TGT GCC CCA GGC GCT GCA ACA GCT GGC CCA GCC CAC
arg gly glu pro val asp glu gln cys ala pro gly ala ala thr ala gly pro ala his
181/61   271/71
GCA GGG CAC CAC GCC TTC TTC CAA GCT GGG TGG CCT GTG GAA GAC GGT CTC GCC GCA TCG
ala gly his his ala phe phe gln ala gly trp pro val glu asp gly leu ala ala ser
241/81
GTC GCC GAT C
val ala asp

SEQ ID NOS.96-97

FIG. 5A'

```
1/1    SEQ ID NO:  98                      31/11
TAC GCC GCG GCG ACG GCG ACG GCG ACG GCG ACG TTG CTG CCG TTC GAG GAG GCG CCG GAG
tyr ala ala ala thr ala thr ala thr ala thr leu leu pro phe glu glu ala pro glu
61/21      SEQ ID NO:  99                              91/31
ATG ACC AGC GCG GGT GGG CTC CTC GAG CAG GCC GCC GCG GTC GAG GAG GCC TCC GAC ACC
met thr ser ala gly gly leu leu glu gln ala ala ala val glu glu ala ser asp thr
121/41                                     151/51
GCC GCG GCG AAC CAG TTG ATG AAC AAT GTG CCC CAG GCG CTG CAA CAG CTG GCC CAG CCC
ala ala ala asn gln leu met asn asn val pro gln ala leu gln gln leu ala gln pro
181/61                                     211/71
ACG CAG GGC ACC ACG CCT TCT TCC AAG CTG GGT GGC CTG TGG AAG ACG GTC TCG CCG CAT
thr gln gly thr thr pro ser ser lys leu gly gly leu trp lys thr val ser pro his
241/81
CGG TCG CCG ATC
arg ser pro ile
```

SEQ ID NOS:98-99

FIG. 5B'

```
1/1    SEQ ID NO:  100                     31/11
ACG CCG CGG CGA CGG CGA CGG CGA CGG CGA CGT TGC TGC CGT TCG AGG AGG CGC CGG AGA
(thr pro arg arg arg arg arg arg arg arg arg cys cys arg ser arg arg arg arg arg)
61/21      SEQ ID NO:  101                  91/31
TGA CCA GCG CGG GTG GGC TCC TCG AGC AGG CCG CCG CGG TCG AGG AGG CCT CCG ACA CCG
OPA(pro ala arg val gly ser ser ser arg pro pro arg ser arg arg pro pro thr pro
121/41      SEQ ID NO:  102                 151/51
CCG CGG CGA ACC AGT TGA TGA ACA ATG TGC CCC AGG CGC TGC AAC AGC TGG CCC AGC CCA
pro arg arg thr ser)OPA OPA(thr met cys pro arg arg cys asn ser trp pro ser pro
181/61     SEQ ID NO:  103                  211/71
CGC AGG GCA CCA CGC CTT CTT CCA AGC TGG GTG GCC TGT GGA AGA CGG TCT CGC CGC ATC
arg arg ala pro arg leu leu pro ser trp val ala cys gly arg arg ser arg arg ile
241/81
GGT CGC CGA TC
gly arg arg)
```

SEQ ID NOS:100-103

FIG. 5C'

ORF predicted by Cole et al. (Nature 393:537-544) and containing seq5A'

1/1 ┌─SEQ ID NO: 104   31/11
tga act gat gat tct gat agc gac caa cct ctt ggg gca aaa cac ccc ggc gat cgc ggt
OPA(thr asp asp ser asp ser asp gln pro leu gly ala lys his pro gly asp arg gly
61/21 ┌──────SEQ ID NO: 105   91/31
caa cga ggc cga ata cgg cga gat gtg ggc cca aga cgc cgc cgc gat gtt tgg cta cgc
gln arg gly arg ile arg arg asp val gly pro arg arg arg asp val trp leu arg
121/41   151/51
cgc ggc gac ggc gac ggc gac ggc gac gtt gct gcc gtt cga gga ggc gcc gga gat gac
arg gly asp gly asp gly asp gly asp val ala ala val arg gly gly ala gly asp asp
181/61   211/71
cag cgc ggg tgg gct cct cga gca ggc cgc cgc ggt cga gga ggc ctc cga cac cgc cgc
gln arg gly trp ala pro arg ala gly arg arg gly arg gly gly leu arg his arg arg
241/81   271/91
ggc gaa cca gtt gat gaa caa tgt gcc cca ggc gct gca aca gct ggc cca gcc cac gca
gly glu pro val asp glu gln cys ala pro gly ala ala thr ala gly pro ala his ala
301/101   331/111
ggg cac cac gcc ttc ttc caa gct ggg tgg cct gtg gaa gac ggt ctc gcc gca tcg gtc
gly his his ala phe phe gln ala gly trp pro val glu asp gly leu ala ala ser val
361/121   391/131
gcc gat cag caa cat ggt gtc gat ggc caa caa cca cat gtc gat gac caa ctc ggg tgt
ala asp gln gln his gly val asp gly gln gln pro his val asp asp gln leu gly cys
421/141   451/151
gtc gat gac caa cac ctt gag ctc gat gtt gaa ggg ctt tgc tcc ggc ggc ggc cgc cca
val asp asp gln his leu glu leu asp val glu gly leu cys ser gly gly gly arg pro
481/161   511/171
ggc cgt gca aac cgc ggc gca aaa cgg ggt ccg ggc gat gag ctc gct ggg cag ctc gct
gly arg ala asn arg gly ala lys arg gly pro gly asp glu leu ala gly gln leu ala
541/181   571/191
ggg ttc ttc ggg tct ggg cgg tgg ggt ggc cgc caa ctt ggg tcg ggc ggc ctc ggt cgg
gly phe phe gly ser gly arg trp gly gly arg gln leu gly ser gly gly leu gly arg
601/201   631/211
ttc gtt gtc ggt gcc gca ggc ctg ggc cgc ggc caa cca ggc agt cac ccc ggc ggc gcg
phe val val gly ala ala gly leu gly arg gly gln pro gly ser his pro gly gly ala
661/221   691/231
ggc gct gcc gct gac cag cct gac cag cgc cgc gga aag agg gcc cgg gca gat gct ggg
gly ala ala ala asp gln pro asp gln arg arg gly lys arg ala arg ala asp ala gly
721/241   751/251
cgg gct gcc ggt ggg gca gat ggg cgc cag ggc cgg tgg tgg gct cag tgg tgt gct gcg
arg ala ala gly gly ala asp gly arg gln gly arg trp trp ala gln trp cys ala ala
781/261   811/271
tgt tcc gcc gcg acc cta tgt gat gcc gca ttc tcc ggc ggc cgg cta gga gag ggg gcg
cys ser ala ala thr leu cys asp ala ala phe ser gly gly arg leu gly glu gly ala
841/281
cag act gtc gtt att tga
gln thr val val ile)OPA

SEQ ID NOS:104-105

FIG. 5D'

sequence Rv1196 predicted by Cole et al. (Nature 393:537-544) and capable
of encoding an ORF fused with Seq5A'

1/1 ┌─SEQ ID NO: 106   31/11
atg gtg gat ttc ggg gcg tta cca ccg gag atc aac tcc gcg agg atg tac gcc ggc ccg
Met val asp phe gly ala leu pro pro glu ile asn ser ala arg met tyr ala gly pro
61/21 ┌─SEQ ID NO: 107   91/31
ggt tcg gcc tcg ctg gtg gcc gcg gct cag atg tgg gac agc gtg gcg agt gac ctg ttt
gly ser ala ser leu val ala ala ala gln met trp asp ser val ala ser asp leu phe
121/41   151/51
tcg gcc gcg tcg gcg ttt cag tcg gtg gtc tgg ggt ctg acg gtg ggg tcg tgg ata ggt
ser ala ala ser ala phe gln ser val val trp gly leu thr val gly ser trp ile gly
181/61   211/71
tcg tcg gcg ggt ctg atg gtg gcg gcg gcc tcg ccg tat gtg gcg tgg atg agc gtc acc
ser ser ala gly leu met val ala
ala ala ser pro tyr val ala trp met ser val thr
241/81   271/91
gcg ggg cag gcc gag ctg acc gcc gcc cag gtc cgg gtt gct gcg gcg gcc tac gag acg
ala gly gln ala glu leu thr ala ala gln val arg val ala ala ala ala tyr glu thr
301/101   331/111
gcg tat ggg ctg acg gtg ccc ccg ccg gtg atc gcc gag aac cgt gct gaa ctg atg att
ala tyr gly leu thr val pro pro pro val ile ala glu asn arg ala glu leu met ile
361/121   391/131
ctg ata gcg acc aac ctc ttg ggg caa aac acc ccg gcg atc gcg gtc aac gag gcc gaa
leu ile ala thr asn leu leu gly gln asn thr pro ala ile ala val asn glu ala glu
421/141   451/151
tac ggc gag atg tgg gcc caa gac gcc gcc gcg atg ttt ggc tac gcc gcg gcg acg gcg
tyr gly glu met trp ala gln asp ala ala ala met phe gly tyr ala ala ala thr ala
481/161   511/171
acg gcg acg gcg acg ttg ctg ccg ttc gag gag gcg ccg gag atg acc agc gcg ggt ggg
thr ala thr ala thr leu leu pro phe glu glu ala pro glu met thr ser ala gly gly
541/181   571/191
ctc ctc gag cag gcc gcc gcg gtc gag gag gcc tcc gac acc gcc gcg gcg aac cag ttg
leu leu glu gln ala ala ala val glu glu ala ser asp thr ala ala ala asn gln leu
601/201   631/211
atg aac aat gtg ccc cag gcg ctg caa cag ctg gcc cag ccc acg cag ggc acc acg cct
met asn asn val pro gln ala leu gln gln leu ala gln pro thr gln gly thr thr pro
661/221   691/231
tct tcc aag ctg ggt ggc ctg tgg aag acg gtc tcg ccg cat cgg tcg ccg atc agc aac
ser ser lys leu gly gly leu trp lys thr val ser pro his arg ser pro ile ser asn
721/241   751/251
atg gtg tcg atg gcc aac aac cac atg tcg atg acc aac tcg ggt gtg tcg atg acc aac
met val ser met ala asn asn his met ser met thr asn ser gly val ser met thr asn
781/261   811/271
acc ttg agc tcg atg ttg aag ggc ttt gct ccg gcg gcg gcc gcc cag gcc gtg caa acc
thr leu ser ser met leu lys gly phe ala pro ala ala ala ala gln ala val gln thr
841/281   871/291
gcg gcg caa aac ggg gtc cgg gcg atg agc tcg ctg ggc agc tcg ctg ggt tct tcg ggt
ala ala gln asn gly val arg ala met ser ser leu gly ser ser leu gly ser ser gly
901/301   931/311
ctg ggc ggt ggg gtg gcc gcc aac ttg ggt cgg gcg gcc tcg gtc ggt tcg ttg tcg gtg
leu gly gly gly val ala ala asn leu gly arg ala ala ser val gly ser leu ser val
961/321   991/331
ccg cag gcc tgg gcc gcg gcc aac cag gca gtc acc ccg gcg gcg cgg gcg ctg ccg ctg
pro gln ala trp ala ala ala asn gln ala val thr pro ala ala arg ala leu pro leu
1021/341   1051/351
acc agc ctg acc agc gcc gcg gaa aga ggg ccc ggg cag atg ctg ggc ggg ctg ccg gtg
thr ser leu thr ser ala ala glu arg gly pro gly gln met leu gly gly leu pro val
1081/361   1111/371
ggg cag atg ggc gcc agg gcc ggt ggt ggg ctc agt ggt gtg ctg cgt gtt ccg ccg cga
gly gln met gly ala arg ala gly gly gly leu ser gly val leu arg val pro pro arg
1141/381   1171/391
ccc tat gtg atg ccg cat tct ccg gcg gcc ggc tag
pro tyr val met pro his ser pro ala ala gly)AMB

SEQ ID NOS:106-107

FIG. 5E'

Seq 5P: ORF according to Cole et al. (Nature 393:537-544) and containing the sequence Rv1196

1/1 ___ SEQ ID NO: 108
tag gga cac gta atg gtg gat ttc ggg gcg
AMB(gly his val met val asp phe gly ala
61/21 ___ SEQ ID NO: 109
tac gcc ggc ccg ggt tcg gcc tcg ctg gtg
tyr ala gly pro gly ser ala ser leu val
121/41
agt gac ctg ttt tcg gcc gcg tcg gcg ttt
ser asp leu phe ser ala ala ser ala phe
181/61
tcg tgg ata ggt tcg tcg gcg ggt ctg atg
ser trp ile gly ser ser ala gly leu met
241/81
atg agc gtc acc gcg ggg cag gcc gag ctg
met ser val thr ala gly gln ala glu leu
301/101
gcc tac gag acg gcg tat ggg ctg acg gtg
ala tyr glu thr ala tyr gly leu thr val
361/121
gaa ctg atg att ctg ata gcg acc aac ctc
glu leu met ile leu ile ala thr asn leu
421/141
aac gag gcc gaa tac ggc gag atg tgg gcc
asn glu ala glu tyr gly glu met trp ala
481/161
gcg gcg acg gcg acg gcg acg gcg acg ttg
ala ala thr ala thr ala thr ala thr leu
541/181
agc gcg ggt ggg ctc ctc gag cag gcc gcc
ser ala gly gly leu leu glu gln ala ala
601/201
gcg aac cag ttg atg aac aat gtg ccc cag
ala asn gln leu met asn asn val pro gln
661/221
ggc acc acg cct tct tcc aag ctg ggt ggc
gly thr thr pro ser ser lys leu gly gly
721/241
ccg atc agc aac atg gtg tcg atg gcc aac
pro ile ser asn met val ser met ala asn
781/261
tcg atg acc aac acc ttg agc tcg atg ttg
ser met thr asn thr leu ser ser met leu
841/281
gcc gtg caa acc gcg gcg caa aac ggg gtc
ala val gln thr ala ala gln asn gly val
901/301
ggt tct tcg ggt ctg ggc ggt ggg gtg gcc
gly ser ser gly leu gly gly gly val ala
961/321
tcg ttg tcg gtg ccg cag gcc tgg gcc gcg
ser leu ser val pro gln ala trp ala ala
1021/341
gcg ctg ccg ctg acc agc ctg acc agc gcc
ala leu pro leu thr ser leu thr ser ala
1081/361 1111/371
ggg ctg ccg gtg ggg cag atg ggc gcc agg
gly leu pro val gly gln met gly ala arg
1141/381
gtt ccg ccg cga ccc tat gtg atg ccg cat
val pro pro arg pro tyr val met pro his

31/11
tta cca ccg gag atc aac tcc gcg agg atg
leu pro pro glu ile asn ser ala arg met
91/31
gcc gcg gct cag atg tgg gac agc gtg gcg
ala ala ala gln met trp asp ser val ala
151/51
cag tcg gtg gtc tgg ggt ctg acg gtg ggg
gln ser val val trp gly leu thr val gly
211/71
gtg gcg gcg gcc tcg ccg tat gtg gcg tgg
val ala ala ala ser pro tyr val ala trp
271/91
acc gcc gcc cag gtc cgg gtt gct gcg gcg
thr ala ala gln val arg val ala ala ala
331/111
ccc ccg ccg gtg atc gcc gag aac cgt gct
pro pro pro val ile ala glu asn arg ala
391/131
ttg ggg caa aac acc ccg gcg atc gcg gtc
leu gly gln asn thr pro ala ile ala val
451/151
caa gac gcc gcc gcg atg ttt ggc tac gcc
gln asp ala ala ala met phe gly tyr ala
511/171
ctg ccg ttc gag gag gcg ccg gag atg acc
leu pro phe glu glu ala pro glu met thr
571/191
gcg gtc gag gag gcc tcc gac acc gcc gcg
ala val glu glu ala ser asp thr ala ala
631/211
gcg ctg caa cag ctg gcc cag ccc acg cag
ala leu gln gln leu ala gln pro thr gln
691/231
ctg tgg aag acg gtc tcg ccg cat cgg tcg
leu trp lys thr val ser pro his arg ser
751/251
aac cac atg tcg atg acc aac tcg ggt gtg
asn his met ser met thr asn ser gly val
811/271
aag ggc ttt gct ccg gcg gcg gcc gcc cag
lys gly phe ala pro ala ala ala ala gln
871/291
cgg gcg atg agc tcg ctg ggc agc tcg ctg
arg ala met ser ser leu gly ser ser leu
931/311
gcc aac ttg ggt cgg gcg gcc tcg gtc ggt
ala asn leu gly arg ala ala ser val gly
991/331
gcc aac cag gca gtc acc ccg gcg gcg cgg
ala asn gln ala val thr pro ala ala arg
1051/351
gcg gaa aga ggg ccc ggg cag atg ctg ggc
ala glu arg gly pro gly gln met leu gly
1171/391
gcc ggt ggt ggg ctc agt ggt gtg ctg cgt
ala gly gly gly leu ser gly val leu arg
tct ccg gcg gcc ggc tag
ser pro ala ala gly)AMB

SEQ ID NOS:108-109

FIG. 5F'

SEQ ID NO: 110   31/11
```
GGA TCC TGA TGC AAG TGG TCC GGG ATT TGT   CGG CAG CCA CCG CGG TCC CGT CGA CCA ACG
gly ser OPA(cys lys trp ser gly ile cys   arg gln pro arg arg ser arg arg pro thr
61/21        SEQ ID NO:   111              91/31
TTG GTG CAT CCG GGC TGC GAG CAT GCA CGC   ACC GAC CAG CGC GGC GAG CGC GGC TAG CTG
leu val his pro gly cys glu his ala arg   thr asp gln arg gly glu arg gly)AMB(leu
121/41                                     151/51              SEQ ID NO:   112
CTT GCC CAC TGT TCC TCC CTG CCG GCA CCA   TGT GCG ACA AGC TTA AGC GCA GCA GTA CCG
leu ala his cys ser ser leu pro ala pro   cys ala thr ser leu ser ala ala val pro
181/61                                     211/71
GCG GTG CCT GGG CAT CCA GCA AAA CGG GGA   GCT CAA GAA CGA TTC ATG AAC GAG GGG TCG
ala val pro gly his pro ala lys arg gly   ala gln glu arg phe met asn glu gly sec
241/81                                     271/91
TCA CCA ACG TCG AAA CCG ACG GTT GCC AGC   CGG CCC ACG ATA TTG CGT GCT CGA GGG TCC
ser pro thr ser lys pro thr val ala ser   arg pro thr ile leu arg ala arg gly ser
301/101                                    331/111
GCT GTA CCC TCA CCG AAC GTG AGT CCC ACA   CCG CGG AGG CGG GCG ACT CTG GCG TCG TTA
ala val pro ser pro asn val ser pro thr   pro arg arg arg ala thr leu ala ser leu
361/121                                    391/131
GCA GCC GAG CTC AAG GTG TCC CGC ACC ACT   GTC TCG AAT GCT TTT AAC CGA CCG GAT CCA
ala ala glu leu lys val ser arg thr thr   val ser asn ala phe asn arg pro asp pro
421/141
GAA GGA GAA GAT C
glu gly glu asp)
```

SEQ ID NOS:110-112

FIG. 6A

SEQ ID NO: 113   32/11
```
GAT CCT GAT GCA AGT GGT CCG GGA TTT GTC   GGC AGC CAC GGC GGT CCC GTC GAC CAA CGT
(asp pro asp ala ser gly pro gly phe val  gly ser his gly gly pro val asp gln arg
62/21        SEQ ID NO:   114             92/31
TGG TGC ATC CGG GCT GCG AGC ATG CAC GCA   CCG ACC AGC GCG GCG AGC GCG GCT AGC TGC
trp cys ile arg ala ala ser met his ala   pro thr ser ala ala ser ala ala ser cys
122/41                                     152/51
TTG CCC ACT GTT CCT CCC TGC CGG CAC CAT   GTG CGA CAA GCT TAA GCG CAG CAG TAC CGG
leu pro thr val pro pro cys arg his his   val arg gln ala)OCH(ala gln gln tyr arg
182/61                                     212/71              SEQ ID NO:   115
CGG TGC CTG GGC ATC CAG CAA AAC GGG GAG   CTC AAG AAC GAT TCA TGA ACG AGG GGT CGT
arg cys leu gly ile gln gln asn gly glu   leu lys asn asp ser)OPA(thr arg gly arg
242/81                                     272/91 SEQ ID NO:   116
CAC CAA CGT CGA AAC CGA CGG TTG CCA GCC   GGC CCA CGA TAT TGC GTG CTC GAG GGT CCG
his gln arg arg asn arg arg leu pro ala   gly pro arg tyr cys val leu glu gly pro
302/101                                    332/111
CTG TAC CCT CAC CGA ACG TGA GTC CCA CAC   CGC GGA GGC GGG CGA CTC TGG CGT CGT TAG
leu tyr pro his arg thr)OPA(val pro his   arg gly gly gly arg leu trp arg arg)AMB
362/121      SEQ ID NO:   117              392/131
CAG CCG AGC TCA AGG TGT CCC GCA CCA CTG   TCT CGA ATG CTT TTA ACC GAC CGG ATC CAG
(gln pro ser ser arg cys pro ala pro leu  ser arg met leu leu thr asp arg ile gln
422/141      SEQ ID NO:   118
AAG GAG AAG ATC
lys glu lys ile)
```

SEQ ID NOS:113-118

FIG. 6B

SEQ ID NO: 119   33/11
ATC CTG ATG CAA GTG GTC CGG GAT TTG TCG GCA GCC ACG GCG GTC CCG TCG ACC AAC GTT
(ile leu met gln val val arg asp leu ser ala ala thr ala val pro ser thr asn val
63/21   SEQ ID NO: 120   93/31
GGT GCA TCC GGG CTG CGA GCA TGC ACG CAC CGA CCA GCG CGG CGA GCG CGG CTA GCT GCT
gly ala ser gly leu arg ala cys thr his arg pro ala arg arg ala arg leu ala ala
123/41   153/51
TGC CCA CTG TTC CTC CCT GCC GGC ACC ATG TGC GAC AAG CTT AAG CGC AGC AGT ACC GGC
cys pro leu phe leu pro ala gly thr met cys asp lys leu lys arg scr ser thr gly
183/61   213/71
GGT GCC TGG GCA TCC AGC AAA ACG GGG AGC TCA AGA ACG ATT CAT GAA CGA GGG GTC GTC
gly ala trp ala ser ser lys thr gly ser ser arg thr ile his glu arg gly val val
243/81   273/91
ACC AAC GTC GAA ACC GAC GGT TGC CAG CCG GCC CAC GAT ATT GCG TGC TCG AGG GTC CGC
thr asn val glu thr asp gly cys gln pro ala his asp ile ala cys ser arg val arg
303/101   333/111
TGT ACC CTC ACC GAA CGT GAG TCC CAC ACC GCG GAG GCG GGC GAC TCT GGC GTC GTT AGC
cys thr leu thr glu arg glu ser his thr ala glu ala gly asp ser gly val val ser
363/121   393/131
AGC CGA GCT CAA GGT GTC CCG CAC CAC TGT CTC GAA TGC TTT TAA CCG ACC GGA TCC AGA
ser arg ala gln gly val pro his his cys leu glu cys phe)OCH(pro thr gly ser arg
423/141   SEQ ID NO: 121
AGG AGA AGA TC
arg arg arg)

SEQ ID NOS:119-121

FIG. 6C

SEQ ID NO: 122   31/11
CCG TCG GCA ACT TGG CCG CTG AGG TCG GCT TGA TCC CTG GGC CGA GGC GGG TCA GCC AAT
(pro ser ala thr trp pro leu arg ser ala)OPA(ser leu gly arg gly gly ser ala asn
61/21   SEQ ID NO: 123   91/31   SEQ ID NO: 124
AGC GGC TCC ATC GGC TTT GCT GGT AGC GGT TCG GCG GGA AGC TAG CGG CGA CGT TGT CGG
ser gly ser ile gly phe ala gly ser gly ser ala gly ser)AMB(arg arg arg cys arg
121/41   151/51   SEQ ID NO: 125
TGG CCG GTG ATA TAT TCG GTC AGA CGG GTA TGG CGG CGG CTG AGG TGA TCT GCG ACA CGC
trp pro val ile tyr trp val arg arg val trp arg arg leu arg)OPA(ser ala thr arg
181/61   211/71   SEQ ID NO: 126
CGC CGC GGT GCT CGA GCC AGG CTT ACG ACC AGG GAA TTT CGA AAA TGT TAT TCA GAA CAT
arg arg gly ala arg ala arg leu thr thr arg glu phe arg lys cys tyr ser glu his
241/81   271/91
CTT GTA TCT CTT CTC CGT GCC ACC CCC TAG GTG TAG TGT TTT CGA GTA CCG GCA GAT CCC
leu val ser leu leu arg ala thr pro)AMB val AMB(cys phe arg val pro ala asp pro
301/101   SEQ ID NO: 127
AGG TTC ACC AGG TCT CAC CAG ATC
arg phe thr arg ser his gln ile)

SEQ ID NOS:122-127

FIG. 7A

SEQ ID NO: 128   32/11
CGT CGG CAA CTT GGC CGC TGA GGT CGG CTT GAT CCC TGG GCC GAG GCG GGT CAG CCA ATA
(arg arg gln leu gly arg)OPA(gly arg leu asp pro trp ala glu ala gly gln pro ile
62/21 SEQ ID NO: 129   92/31 SEQ ID NO: 130
GCG GCT CCA TCG GCT TTG CTG GTA GCG GTT CGG CGG GAA GCT AGC GGC GAC GTT GTC GGT
ala ala pro ser ala leu leu val ala val arg arg glu ala ser gly asp val val gly
122/41   152/51
GGC CGG TGA TAT ATT GGG TCA GAC GGG TAT GGC GGC GGC TGA GGT GAT CTG CGA CAC GCC
gly arg)OPA(tyr ile gly ser asp gly tyr gly gly gly)OPA(gly asp leu arg his ala
182/61   SEQ ID NO: 131   212/71   SEQ ID NO: 132
GCC GCG GTG CTC GAG CCA GGC TTA CGA CCA GGG AAT TTC GAA AAT GTT ATT CAG AAC ATC
ala ala val leu glu pro gly leu arg pro gly asn phe glu asn val ile gln asn ile
242/81   272/91
TTG TAT CTC TTC TCC GTG CCA CCC CCT AGG TGT AGT GTT TTC GAG TAC CGG CAG ATC CCA
leu tyr leu phe ser val pro pro pro arg cys ser val phe glu tyr arg gln ile pro
302/101
GGT TCA CCA GGT CTC ACC AGA TC
gly ser pro gly leu thr arg)

SEQ ID NOS:128-132

FIG. 7B

SEQ ID NO: 133   33/11
GTC GGC AAC TTG GCC GCT GAG GTC GGC TTG ATC CCT GGG CCG AGG CGG GTC AGC CAA TAG
(val gly asn leu ala ala glu val gly leu ile pro gly pro arg arg val ser gln)AMB
63/21   SEQ ID NO: 134   93/31
CGG CTC CAT CGG CTT TGC TGG TAG CGG TTC GGC GGG AAG CTA GCG GCG ACG TTG TCG GTG
(arg leu his arg leu cys trp)AMB(arg phe gly gly lys leu ala ala thr leu ser val
123/41   SEQ ID NO: 135   153/51   SEQ ID NO: 136
GCC GGT GAT ATA TTG GGT CAG ACG GGT ATG GCG GCG GCT GAG GTG ATC TGC GAC ACG CCG
ala gly asp ile leu gly gln thr gly met ala ala ala glu val ile cys asp thr pro
183/61   213/71
CCG CGG TGC TCG AGC CAG GCT TAC GAC CAG GGA ATT TCG AAA ATG TTA TTC AGA ACA TCT
pro arg cys ser ser gln ala tyr asp gln gly ile ser lys met leu phe arg thr ser
243/81   273/91
TGT ATC TCT TCT CCG TGC CAC CCC CTA GGT GTA GTG TTT TCG AGT ACC GGC AGA TCC CAG
cys ile ser ser pro cys his pro leu gly val val phe ser ser thr gly arg ser gln
303/101
GTT CAC CAG GTC TCA CCA GAT C
val his gln val ser pro asp)

SEQ ID NOS:133-136

FIG. 7C

SEQ ID NO: 137   31/11
CTT TGC GTG ATG TCC AAT GGC GAA AAC GAC GCC TTG TCA TCG CAA TCG TCA GCA CCG GCC
(leu cys val met ser asn gly glu asn asp ala leu ser ser gln ser ser ala pro ala)
61/21   SEQ ID NO: 138   91/31
TAG TTT TCG CGA TGA CGC TCG TTC TGA CCG GAC TTG TGA ACG GGT TTC GGG TCG AGG CCG
AMB phe ser arg OPA arg ser phe OPA pro asp leu OPA(thr gly phe gly ser arg pro
121/41   151/51   SEQ ID NO: 272
AGC GAA CCG TCG ATT CCA TGG GTG TCG ACG CAT TCG TGG TCA AGG CCG GCG CGG CAG GAC
ser glu pro ser ile pro trp val ser thr his ser trp ser arg pro ala arg gln asp
181/61   211/71
CGT TCC TGG GTT CGA CAC CAT TCG CCC AAA TCG ACC TGC CCC AGG TTG CTC GTG CGC CTG
arg ser trp val arg his his ser pro lys ser thr cys pro arg leu leu val arg leu
241/81   271/91
GCG TCT TGG CTG CCG CCC CAC TAG CGA CTG CGC CGT CGA CGA TCC GGC AGG GCA CGT CAG
ala ser trp leu pro pro his)AMB(arg leu arg arg arg arg ser gly arg ala arg gln
301/101   SEQ ID NO: 273   331/111
CGC GAA ACG TCA CCG CGT TCG GGG CAC CAG AGC ACG GAC CCG GCA TGC CGC GGG TCT CGG
arg glu thr ser pro arg ser gly his gln ser thr asp pro ala cys arg gly ser arg
361/121   391/131
ACG GTC GGG CGC CAT CGA CGC CGG ACG AGG TCG CGG TGT CGA GCA CGC TGG GCC GAA ACC
thr val gly arg his arg arg arg thr arg ser arg cys arg ala arg trp ala glu thr
421/141
TCG GCG ACG ATC
ser ala thr ile)

SEQ ID NOS:137-138,272-273

FIG. 8A

SEQ ID NO: 139   32/11
TTT GCG TGA TGT CCA ATG GCG AAA ACG ACG CCT TGT CAT CGC AAT CGT CAG CAC CGG CCT
phe ala OPA(cys pro met ala lys thr thr pro cys his arg asn arg gln his arg pro
62/21   SEQ ID NO: 140   92/31
AGT TTT CGC GAT GAC GCT CGT TCT GAC CGG ACT TGT GAA CGG GTT TCG GGT CGA GGC CGA
ser phe arg asp asp ala arg ser asp arg thr cys glu arg val ser gly arg gly arg
122/41   152/51
GCG AAC CGT CGA TTC CAT GGG TGT CGA CGC ATT CGT GGT CAA GGC CGG CGC GGC AGG ACC
ala asn arg arg phe his gly cys arg arg ile arg gly gln gly arg arg gly arg thr
182/61   212/71
GTT CCT GGG TTC GAC ACC ATT CGC CCA AAT CGA CCT GCC CCA GGT TGC TCG TGC GCC TGG
val pro gly phe asp thr ile arg pro asn arg pro ala pro gly cys ser cys ala trp
242/81   272/91
CGT CTT GGC TGC CGC CCC ACT AGC GAC TGC GCC GTC GAC GAT CCG GCA GGG CAC GTC AGC
arg leu gly cys arg pro thr ser asp cys ala val asp asp pro ala gly his val ser
302/101   332/111
GCG AAA CGT CAC CGC GTT CGG GGC ACC AGA GCA CGG ACC CGG CAT GCC GCG GGT CTC GGA
ala lys arg his arg val arg gly thr arg ala arg thr arg his ala ala gly leu gly
362/121   392/131
CGG TCG GGC GCC ATC GAC GCC GGA CGA GGT CGC GGT GTC GAG CAC GCT GGG CCG AAA CCT
arg ser gly ala ile asp ala gly arg gly arg gly val glu his ala gly pro lys pro
422/141
CGG CGA CGA TC
arg arg arg)

SEQ ID NOS:139-140

FIG. 8B

SEQ ID NO: 141   33/11
TTG CGT GAT GTC CAA TGG CGA AAA CGA CGC CTT GTC ATC GCA ATC GTC AGC ACC GGC CTA
(leu arg asp val gln trp arg lys arg arg leu val ile ala ile val ser thr gly leu
63/21 SEQ ID NO: 142   93/31
GTT TTC GCG ATG ACG CTC GTT CTG ACC GGA CTT GTG AAC GGG TTT CGG GTC GAG GCC GAG
val phe ala met thr leu val leu thr gly leu val asn gly phe arg val glu ala glu
123/41   153/51
CGA ACC GTC GAT TCC ATG GGT GTC GAC GCA TTC GTG GTC AAG GCC GGC GCG GCA GGA CCG
arg thr val asp ser met gly val asp ala phe val val lys ala gly ala ala gly pro
183/61   213/71
TTC CTG GGT TCG ACA CCA TTC GCC CAA ATC GAC CTG CCC CAG GTT GCT CGT GCG CCT GGC
phe leu gly ser thr pro phe ala gln ile asp leu pro gln val ala arg ala pro gly
243/81   273/91
GTC TTG GCT GCC GCC CCA CTA GCG ACT GCG CCG TCG ACG ATC CGG CAG GGC ACG TCA GCG
val leu ala ala ala pro leu ala thr ala pro ser thr ile arg gln gly thr ser ala
303/101   333/111
CGA AAC GTC ACC GCG TTC GGG GCA CCA GAG CAC GGA CCC GGC ATG CCG CGG GTC TCG GAC
arg asn val thr ala phe gly ala pro glu his gly pro gly met pro arg val ser asp
363/121   393/131
GGT CGG GCG CCA TCG ACG CCG GAC GAG GTC GCG GTG TCG AGC ACG CTG GGC CGA AAC CTC
gly arg ala pro ser thr pro asp glu val ala val ser ser thr leu gly arg asn leu
423/141
GGC GAC GAT C
gly asp asp)

SEQ ID NOS:141-142

FIG. 8C

part of the nucleotide sequence of seq8A

1/1 SEQ ID NO: 143   31/11
CAG GTT GCT CGT GCG CCT GGC GTC TTG GCT GCC GCC CCA CTA GCG ACT GCG CCG TCG ACG
(gln val ala arg ala pro gly val leu ala ala ala pro leu ala thr ala pro ser thr
61/21 SEQ ID NO: 144   91/31
ATC CGG CAG GGC ACG TCA GCG CGA AAC GTC ACC GCG TTC GGG GCA CCA GAG CAC GGA CCC
ile arg gln gly thr ser ala arg asn val thr ala phe gly ala pro glu his gly pro
121/41   151/51
GGC ATG CCG CGG GTC TCG GAC GGT CGG GCG CCA TCG ACG CCG GAC GAG GTC GCG GTG TCG
gly met pro arg val ser asp gly arg ala pro ser thr pro asp glu val ala val ser
181/61
AGC ACG CTG GGC CGA AAC CTC GGC GAC GAT C
ser thr leu gly arg asn leu gly asp asp)

SEQ ID NOS.143-144

FIG. 8A'

1/1 — SEQ ID NO: 145   31/11
AGG TTG CTC GTG CGC CTG GCG TCT TGG CTG CCG CCC CAC TAG CGA CTG CGC CGT CGA CGA
(arg leu leu val arg leu ala ser trp leu pro pro his)AMB(arg leu arg arg arg arg
61/21 — SEQ ID NO: 146   91/31   — SEQ ID NO: 147
TCC GGC AGG GCA CGT CAG CGC GAA ACG TCA CCG CGT TCG GGG CAC CAG AGC ACG GAC CCG
ser gly arg ala arg gln arg glu thr ser pro arg ser gly his gln ser thr asp pro
121/41   151/51
GCA TGC CGC GGG TCT CGG ACG GTC GGG CGC CAT CGA CGC CGG ACG AGG TCG CGG TGT CGA
ala cys arg gly ser arg thr val gly arg his arg arg arg thr arg ser arg cys arg
181/61
GCA CGC TGG GCC GAA ACC TCG GCG ACG ATC
ala arg trp ala glu thr ser ala thr ile)

SEQ ID NOS:145-147

FIG. 8B'

Seq8C

1/1 — SEQ ID NO: 148   31/11
CCA GGT TGC TCG TGC GCC TGG CGT CTT GGC TGC CGC CCC ACT AGC GAC TGC GCC GTC GAC
(pro gly cys ser cys ala trp arg leu gly cys arg pro thr ser asp cys ala val asp
61/21 — SEQ ID NO: 149   91/31
GAT CCG GCA GGG CAC GTC AGC GCG AAA CGT CAC CGC GTT CGG GGC ACC AGA GCA CGG ACC
asp pro ala gly his val ser ala lys arg his arg val arg gly thr arg ala arg thr
121/41   151/51
CGG CAT GCC GCG GGT CTC GGA CGG TCG GGC GCC ATC GAC GCC GGA CGA GGT CGC GGT GTC
arg his ala ala gly leu gly arg ser gly ala ile asp ala gly arg gly arg gly val
181/61
GAG CAC GCT GGG CCG AAA CCT CGG CGA CGA TC
glu his ala gly pro lys pro arg arg arg)

SEQ ID NOS:148-149

FIG. 8C'

sequence Rv2563 predicted by Cole et al. (Nature 393:537-544) and
containing seq8A'

atg
met
121/41 ___ SEQ ID NO: 150   151/51
ctt ttt gcg gct ttg cgt gat gtc caa tgg cga aaa cga cgc ctt gtc atc gca atc gtc
(leu phe ala ala leu arg asp val gln trp arg lys arg arg leu val ile ala ile val
181/61 ___ SEQ ID NO: 151   211/71
agc acc ggc cta gtt ttc gcg atg acg ctc gtt ctg acc gga ctt gtg aac ggg ttt cgg
ser thr gly leu val phe ala met thr leu val leu thr gly leu val asn gly phe arg
241/81   271/91
gtc gag gcc gag cga acc gtc gat tcc atg ggt gtc gac gca ttc gtg gtc aag gcc ggc
val glu ala glu arg thr val asp ser met gly val asp ala phe val val lys ala gly
301/101   331/111
gcg gca gga ccg ttc ctg ggt tcg aca cca ttc gcc caa atc gac ctg ccc cag gtt gct
ala ala gly pro phe leu gly ser thr pro phe ala gln ile asp leu pro gln val ala
361/121   391/131
cgt gcg cct ggc gtc ttg gct gcc gcc cca cta gcg act gcg ccg tcg acg atc cgg cag
arg ala pro gly val leu ala ala ala pro leu ala thr ala pro ser thr ile arg gln
421/141   451/151
ggc acg tca gcg cga aac gtc acc gcg ttc ggg gca cca gag cac gga ccc ggc atg ccg
gly thr ser ala arg asn val thr ala phe gly ala pro glu his gly pro gly met pro
481/161   511/171
cgg gtc tcg gac ggt cgg gcg cca tcg acg ccg gac gag gtc gcg gtg tcg agc acg ctg
arg val ser asp gly arg ala pro ser thr pro asp glu val ala val ser ser thr leu
541/181   571/191
ggc cga aac ctc ggc gac gat ctg caa gtg ggt gcg cgc act ttg cgg atc gtc ggc atc
gly arg asn leu gly asp asp leu gln val gly ala arg thr leu arg ile val gly ile
601/201   631/211
gtg ccc gag tca acc gcg ctg gca aag att ccc aac atc ttc ctg acc acc gaa ggc cta
val pro glu ser thr ala leu ala lys ile pro asn ile phe leu thr thr glu gly leu
661/221   691/231
cag cag ttg gca tac aac gga cag ccg aca atc agt tcg atc ggg atc gac ggg atg ccc
gln gln leu ala tyr asn gly gln pro thr ile ser ser ile gly ile asp gly met pro
721/241   751/251
cga cag ctc ccg gac ggc tat cag acc gtc aat cga gcg gat gct gtc agc gat ctg atg
arg gln leu pro asp gly tyr gln thr val asn arg ala asp ala val ser asp leu met
781/261   811/271
cgc ccg ttg aag gtc gcg gtg gat gcg atc acg gtt gtg gcg gtc ttg ctg tgg atc gtt
arg pro leu lys val ala val asp ala ile thr val val ala val leu leu trp ile val
841/281   871/291
gcg gcg ttg atc gtc ggc tcg gtg gtc tac ctc tct gcg ttg gag cgg ctg cgt gac ttt
ala ala leu ile val gly ser val val tyr leu ser ala leu glu arg leu arg asp phe
901/301   931/311
gcg gtg ttc aag gcg atc ggc gtg ccg acg cgc tcg att ctg gcc ggg ctg gcg ctg cag
ala val phe lys ala ile gly val pro thr arg ser ile leu ala gly leu ala leu gln
961/321   991/331
gcg gtc gtc gtc gcg ctg ctc gcg gcg gtg gtt ggc ggc atc ctt tcg ctg ctg ttg gcg
ala val val val ala leu leu ala ala val val gly gly ile leu ser leu leu leu ala
1021/341   1051/351
ccg ttg ttc ccg atg act gtc gtg gta ccc ctg agt gcc ttc gtg gcg cta ccg gcg atc
pro leu phe pro met thr val val val pro leu ser ala phe val ala leu pro ala ile
1081/361   1111/371
gcg act gtg atc ggt ctg ctg gcc agc gtc gca gga ctg cgg cgc gtg gtg gcg atc gat
ala thr val ile gly leu leu ala ser val ala gly leu arg arg val val ala ile asp
1141/381
ccg gca cta gcg ttc gga ggt ccc tag
pro ala leu ala phe gly gly pro)AMB

SEQ ID NOS:150-151

FIG. 8D

ORF predicted by Cole et al. (Nature 393:537-544) and containing Rv2563
1/1 ⌐SEQ ID NO: 152   31/11
tag gtt tca aga agg cct gtg cag gtt tcc gca gcc tgg gcc gcg gcg cca ccg aag agc
AMB(val ser arg arg pro val gln val ser ala ala trp ala ala ala pro pro lys ser
61/21 ⌐SEQ ID NO: 153   91/31
ccg ccg aaa tgg gct aat cgg gtt cgc ttg gct cga tcg ccg atg atc tcg acc gcc acg
pro pro lys trp ala asn arg val arg leu ala arg ser pro met ile ser thr ala thr
121/41   151/51
acc gac ccc ctc acc tcg gtc gaa cct cgg cga acc aac gcg gca acg cca gcc cat gat
thr asp pro leu thr ser val glu pro arg arg thr asn ala ala thr pro ala his asp
181/61   211/71
cat ttg att ggg tcc acg gaa gca ggt agc ttc cgt cgc atg ctt ttt gcg gct ttg cgt
his leu ile gly ser thr glu ala gly ser phe arg arg met leu phe ala ala leu arg
241/81   271/91
gat gtc caa tgg cga aaa cga cgc ctt gtc atc gca atc gtc agc acc ggc cta gtt ttc
asp val gln trp arg lys arg arg leu val ile ala ile val ser thr gly leu val phe
301/101   331/111
gcg atg acg ctc gtt ctg acc gga ctt gtg aac ggg ttt cgg gtc gag gcc gag cga acc
ala met thr leu val leu thr gly leu val asn gly phe arg val glu ala glu arg thr
361/121   391/131
gtc gat tcc atg ggt gtc gac gca ttc gtg gtc aag gcc ggc gcg gca gga ccg ttc ctg
val asp ser met gly val asp ala phe val val lys ala gly ala ala gly pro phe leu
421/141   451/151
ggt tcg aca cca ttc gcc caa atc gac ctg ccc cag gtt gct cgt gcg cct ggc gtc ttg
gly ser thr pro phe ala gln ile asp leu pro gln val ala arg ala pro gly val leu
481/161   511/171
gct gcc gcc cca cta gcg act gcg ccg tcg acg atc cgg cag ggc acg tca gcg cga aac
ala ala ala pro leu ala thr ala pro ser thr ile arg gln gly thr ser ala arg asn
541/181   571/191
gtc acc gcg ttc ggg gca cca gag cac gga ccc ggc atg ccg cgg gtc tcg gac ggt cgg
val thr ala phe gly ala pro glu his gly pro gly met pro arg val ser asp gly arg
601/201   631/211
gcg cca tcg acg ccg gac gag gtc gcg gtg tcg agc acg ctg ggc cga aac ctc ggc gac
ala pro ser thr pro asp glu val ala val ser ser thr leu gly arg asn leu gly asp
661/221   691/231
gat ctg caa gtg ggt gcg cgc act ttg cgg atc gtc ggc atc gtg ccc gag tca acc gcg
asp leu gln val gly ala arg thr leu arg ile val gly ile val pro glu ser thr ala
721/241   751/251
ctg gca aag att ccc aac atc ttc ctg acc acc gaa ggc cta cag cag ttg gca tac aac
leu ala lys ile pro asn ile phe leu thr thr glu gly leu gln gln leu ala tyr asn
781/261   811/271
gga cag ccg aca atc agt tcg atc ggg atc gac ggg atg ccc cga cag ctc ccg gac ggc
gly gln pro thr ile ser ser ile gly ile asp gly met pro arg gln leu pro asp gly
841/281   871/291
tat cag acc gtc aat cga gcg gat gct gtc agc gat ctg atg cgc ccg ttg aag gtc gcg
tyr gln thr val asn arg ala asp ala val ser asp leu met arg pro leu lys val ala
901/301   931/311
gtg gat gcg atc acg gtt gtg gcg gtc ttg ctg tgg atc gtt gcg gcg ttg atc gtc ggc
val asp ala ile thr val val ala val leu leu trp ile val ala ala leu ile val gly
961/321   991/331
tcg gtg gtc tac ctc tct gcg ttg gag cgg ctg cgt gac ttt gcg gtg ttc aag gcg atc
ser val val tyr leu ser ala leu glu arg leu arg asp phe ala val phe lys ala ile
1021/341   1051/351
ggc gtg ccg acg cgc tcg att ctg gcc ggg ctg gcg ctg cag gcg gtc gtc gtc gcg ctg
gly val pro thr arg ser ile leu ala gly leu ala leu gln ala val val val ala leu
1081/361   1111/371
ctc gcg gcg gtg gtt ggc ggc atc ctt tcg ctg ctg ttg gcg ccg ttg ttc ccg atg act
leu ala ala val val gly gly ile leu ser leu leu leu ala pro leu phe pro met thr
1141/381   1171/391
gtc gtg gta ccc ctg agt gcc ttc gtg gcg cta ccg gcg atc gcg act gtg atc ggt ctg
val val val pro leu ser ala phe val ala leu pro ala ile ala thr val ile gly leu
1201/401   1231/411
ctg gcc agc gtc gca gga ctg cgg cgc gtg gtg gcg atc gat ccg gca cta gcg ttc gga
leu ala ser val ala gly leu arg arg val val ala ile asp pro ala leu ala phe gly
1261/421
ggt ccc tag
gly pro) AMB

SEQ ID NOS:152-153

FIG. 8E

sequence of Rv0072 predicted by Cole et al. (Nature 393:537-544) and
exhibiting more than 77% similarity with Seq8D'

1/1 ⎯ SEQ ID NO: 154   31/11
atg ctc ttc gcg gcc ctg cgt gac atg caa tgg aga aag cgc cgc ctg gtc atc acg atc
(Met leu phe ala ala leu arg asp met gln trp arg lys arg arg leu val ile thr ile
61/21 ⎯ SEQ ID NO: 155   91/31
atc agc acc ggg ctg atc ttc ggg atg acg ctt gtt ttg acc gga ctc gcg aac ggc ttc
ile ser thr gly leu ile phe gly met thr leu val leu thr gly leu ala asn gly phe
121/41   151/51
cgg gtg gag gcc cgg cac acc gtc gat tcc atg ggt gtc gat gta ttc gtc gtc aga tcc
arg val glu ala arg his thr val asp ser met gly val asp val phe val val arg ser
181/61   211/71
ggc gct gct gga cct ttt ctg ggt tca ata ccg ttt ccc gat gtt gac ctg gcc cga gtg
gly ala ala gly pro phe leu gly ser ile pro phe pro asp val asp leu ala arg val
241/81   271/91
gcc gct gaa ccc ggt gtc atg gcc gcg gcc ccg ttg ggc agc gtg ggg acg atc atg aaa
ala ala glu pro gly val met ala ala ala pro leu gly ser val gly thr ile met lys
301/101   331/111
gaa ggc acg tcg acg cga aac gtc acg gtc ttc ggc gcg ccc gag cac gga cct ggc atg
glu gly thr ser thr arg asn val thr val phe gly ala pro glu his gly pro gly met
361/121   391/131
cca cgg gtc tca gag ggt cgg tca ccg tcg aaa ccg gac gaa gtc gcg gca tcg agc acg
pro arg val ser glu gly arg ser pro ser lys pro asp glu val ala ala ser ser thr
421/141   451/151
atg ggc cga cac ctc ggt gac act gtc gag gtc ggc gcg cgc aga ttg cgg gtc gtt ggc
met gly arg his leu gly asp thr val glu val gly ala arg arg leu arg val val gly
481/161   511/171
att gtg ccg aat tcc acc gcg ctg gcc aag atc ccc aat gtc ttc ctc acg acc gag ggc
ile val pro asn ser thr ala leu ala lys ile pro asn val phe leu thr thr glu gly
541/181   571/191
tta cag aaa ttg gcg tac aac ggg cag ccg aat atc acg tcc atc ggg atc ata ggt atg
leu gln lys leu ala tyr asn gly gln pro asn ile thr ser ile gly ile ile gly met
601/201   631/211
ccc cga cag ctg ccg gag ggt tac cag act ttc gat cgg gtg ggc gct gtc aat gat ttg
pro arg gln leu pro glu gly tyr gln thr phe asp arg val gly ala val asn asp leu
661/221   691/231
gtg cgc cca ttg aag gtc gca gtg aat tcg atc tcg atc gtg gct gtt ttg ctg tgg att
val arg pro leu lys val ala val asn ser ile ser ile val ala val leu leu trp ile
721/241   751/251
gtg gcg gtg ctg atc gtc ggc tcg gtg gtg tac ctt tcg gct ctt gag cgg cta cgt gac
val ala val leu ile val gly ser val val tyr leu ser ala leu glu arg leu arg asp
781/261   811/271
ttc gcg gtg ttc aag gcg att ggc acg cca acg cgc tcg att atg gcc ggg ctc gca tta
phe ala val phe lys ala ile gly thr pro thr arg ser ile met ala gly leu ala leu
841/281   871/291
cag gcg ctg gtc att gcg ttg ctt gcg gcg gtg gtg ggc gtc gtc ctg gcg cag gtg ttg
gln ala leu val ile ala leu leu ala ala val val gly val val leu ala gln val leu
901/301   931/311
gca cca ctg ttt ccg atg att gtc gcg gta ccc gtc ggt gct tac ctg gcg cta ccg gtg
ala pro leu phe pro met ile val ala val pro val gly ala tyr leu ala leu pro val
961/321   991/331
gcc gcg atc gtc atc ggt ctg ttc gct agt gtt gcc gga ttg aag cgc gtg gtg acg gtc
ala ala ile val ile gly leu phe ala ser val ala gly leu lys arg val val thr val
1021/341
gat ccc gcg cag gcg ttc gga ggt ccc tag
asp pro ala gln ala phe gly gly pro)AMB

SEQ ID NOS:154-155

FIG. 8F

Seq8H : ORF predicted by Cole et al. (Nature 393:537-544) and containing seq8G

1/1 — SEQ ID NO: 156   31/11
tag cct ctg gga atg ctc ttc gcg gcc ctg   cgt gac atg caa tgg aga aag cgc cgc ctg
AMB(pro leu gly met leu phe ala ala leu   arg asp met gln trp arg lys arg arg leu

61/21 — SEQ ID NO: 157   91/31
gtc atc acg atc atc agc acc ggg ctg atc   ttc ggg atg acg ctt gtt ttg acc gga ctc
val ile thr ile ile ser thr gly leu ile   phe gly met thr leu val leu thr gly leu

121/41   151/51
gcg aac ggc ttc cgg gtg gag gcc cgg cac   acc gtc gat tcc atg ggt gtc gat gta ttc
ala asn gly phe arg val glu ala arg his   thr val asp ser met gly val asp val phe

181/61   211/71
gtc gtc aga tcc ggc gct gct gga cct ttt   ctg ggt tca ata ccg ttt ccc gat gtt gac
val val arg ser gly ala ala gly pro phe   leu gly ser ile pro phe pro asp val asp

241/81   271/91
ctg gcc cga gtg gcc gct gaa ccc ggt gtc   atg gcc gcg gcc ccg ttg ggc agc gtg ggg
leu ala arg val ala ala glu pro gly val   met ala ala ala pro leu gly ser val gly

301/101   331/111
acg atc atg aaa gaa ggc acg tcg acg cga   aac gtc acg gtc ttc ggc gcg ccc gag cac
thr ile met lys glu gly thr ser thr arg   asn val thr val phe gly ala pro glu his

361/121   391/131
gga cct ggc atg cca cgg gtc tca gag ggt   cgg tca ccg tcg aaa ccg gac gaa gtc gcg
gly pro gly met pro arg val ser glu gly   arg ser pro ser lys pro asp glu val ala

421/141   451/151
gca tcg agc acg atg ggc cga cac ctc ggt   gac act gtc gag gtc ggc gcg cgc aga ttg
ala ser ser thr met gly arg his leu gly   asp thr val glu val gly ala arg arg leu

481/161   511/171
cgg gtc gtt ggc att gtg ccg aat tcc acc   gcg ctg gcc aag atc ccc aat gtc ttc ctc
arg val val gly ile val pro asn ser thr   ala leu ala lys ile pro asn val phe leu

541/181   571/191
acg acc gag ggc tta cag aaa ttg gcg tac   aac ggg cag ccg aat atc acg tcc atc ggg
thr thr glu gly leu gln lys leu ala tyr   asn gly gln pro asn ile thr ser ile gly

601/201   631/211
atc ata ggt atg ccc cga cag ctg ccg gag   ggt tac cag act ttc gat cgg gtg ggc gct
ile ile gly met pro arg gln leu pro glu   gly tyr gln thr phe asp arg val gly ala

661/221   691/231
gtc aat gat ttg gtg cgc cca ttg aag gtc   gca gtg aat tcg atc tcg atc gtg gct gtt
val asn asp leu val arg pro leu lys val   ala val asn ser ile ser ile val ala val

721/241   751/251
ttg ctg tgg att gtg gcg gtg ctg atc gtc   ggc tcg gtg gtg tac ctt tcg gct ctt gag
leu leu trp ile val ala val leu ile val   gly ser val val tyr leu ser ala leu glu

781/261   811/271
cgg cta cgt gac ttc gcg gtg ttc aag gcg   att ggc acg cca acg cgc tcg att atg gcc
arg leu arg asp phe ala val phe lys ala   ile gly thr pro thr arg ser ile met ala

841/281   871/291
ggg ctc gca tta cag gcg ctg gtc att gcg   ttg ctt gcg gcg gtg gtg ggc gtc gtc ctg
gly leu ala leu gln ala leu val ile ala   leu leu ala ala val val gly val val leu

901/301   931/311
gcg cag gtg ttg gca cca ctg ttt ccg atg   att gtc gcg gta ccc gtc ggt gct tac ctg
ala gln val leu ala pro leu phe pro met   ile val ala val pro val gly ala tyr leu

961/321   991/331
gcg cta ccg gtg gcc gcg atc gtc atc ggt   ctg ttc gct agt gtt gcc gga ttg aag cgc
ala leu pro val ala ala ile val ile gly   leu phe ala ser val ala gly leu lys arg

1021/341   1051/351
gtg gtg acg gtc gat ccc gcg cag gcg ttc   gga ggt ccc tag
val val thr val asp pro ala gln ala phe   gly gly pro)AMB

SEQ ID NOS:156-157

FIG. 8G

SEQ ID NO: 158   31/11
CGA GGC CGA GCG AAC CGT CGA TTC CAT GGG TGT CGA CGC ATT CGT GGT CAA GGC CGG CGC
(arg gly arg ala asn arg arg phe his gly cys arg arg ile arg gly gln gly arg arg
61/21   SEQ ID NO: 159   91/31
GGC AGG ACC GTT CCT GGG TTC GAC ACC ATT CGC CCA AAT CGA CCT GCC CCA GGT TGC TCG
gly arg thr val pro gly phe asp thr ile arg pro asn arg pro ala pro gly cys ser
121/41   151/51
TGC GCC TGG CGT CTT GGC TGC CGC CCC ACT AGC GAC TGC GCC GTC GAC GAT CCG GCA GGG
cys ala trp arg leu gly cys arg pro thr ser asp cys ala val asp asp pro ala gly
181/61   211/71
CAC GTC AGC GCG AAA CGT CAC CGC GTT CGG GGC ACC AGA GCA CGG ACC CGG CAT GCC GCG
his val ser ala lys arg his arg val arg gly thr arg ala arg thr arg his ala ala
241/81   271/91
GGT CTC GGA CGG TCG GGC GCC ATC GAC GCC GGA CGA GGT CGC GGT GTC GAG CAC GCT GGG
gly leu gly arg ser gly ala ile asp ala gly arg gly arg gly val glu his ala gly
301/101
CCG AAA CCT CGG CGA CGA TC
pro lys pro arg arg arg)

SEQ ID NOS:158-159

FIG. 9A

SEQ ID NO: 160   32/11
GAG GCC GAG CGA ACC GTC GAT TCC ATG GGT GTC GAC GCA TTC GTG GTC AAG GCC GGC GCG
(glu ala glu arg thr val asp ser met gly val asp ala phe val val lys ala gly ala
62/21   SEQ ID NO: 161   92/31
GCA GGA CCG TTC CTG GGT TCG ACA CCA TTC GCC CAA ATC GAC CTG CCC GAG GTT GCT CGT
ala gly pro phe leu gly ser thr pro phe ala gln ile asp leu pro gln val ala arg
122/41   152/51
GCG CCT GGC GTC TTG GCT GCC GCC CCA CTA GCG ACT GCG CCG TCG ACG ATC CGG CAG GGC
ala pro gly val leu ala ala ala pro leu ala thr ala pro ser thr ile arg gln gly
182/61   212/71
ACG TCA GCG CGA AAC GTC ACC GCG TTC GGG GCA CCA GAG CAC GGA CCC GGC ATG CCG CGG
thr ser ala arg asn val thr ala phe gly ala pro glu his gly pro gly met pro arg
242/81   272/91
GTC TCG GAC GGT CGG GCG CCA TCG ACG CCG GAC GAG GTC GCG GTG TCG AGC ACG CTG GGC
val ser asp gly arg ala pro ser thr pro asp glu val ala val ser ser thr leu gly
302/101
CGA AAC CTC GGC GAC GAT C
arg asn leu gly asp asp)

SEQ ID NOS:160-161

FIG. 9B

SEQ ID NO: 162   33/11
AGG CCG AGC GAA CCG TCG ATT CCA TGG GTG TCG ACG CAT TCG TGG TCA AGG CCG GCG CGG
(arg pro ser glu pro ser ile pro trp val ser thr his ser trp ser arg pro ala arg
63/21   SEQ ID NO: 163   93/31
CAG GAC CGT TCC TGG GTT CGA CAC CAT TCG CCC AAA TCG ACC TGC CCC AGG TTG CTC GTG
gln asp arg ser trp val arg his his ser pro lys ser thr cys pro arg leu leu val
123/41   153/51
CGC CTG GCG TCT TGG CTG CCG CCC CAC TAG CGA CTG CGC CGT CGA CGA TCC GGC AGG GCA
arg leu ala ser trp leu pro pro his)AMB(arg leu arg arg arg arg ser gly arg ala
183/61   SEQ ID NO: 164   213/71
CGT CAG CGC GAA ACG TCA CCG CGT TCG GGG CAC CAG AGC ACG GAC CCG GCA TGC CGC GGG
arg gln arg glu thr ser pro arg ser gly his gln ser thr asp pro ala cys arg gly
243/81   273/91
TCT CGG ACG GTC GGG CGC CAT CGA CGC CGG ACG AGG TCG CGG TGT CGA GCA CGC TGG GCC
ser arg thr val gly arg his arg arg arg thr arg ser arg cys arg ala arg trp ala
303/101
GAA ACC TCG GCG ACG ATC
glu thr.ser ala thr ile)

SEQ ID NOS:162-164

FIG. 9C

SEQ ID NO: 165   31/11
TTA ACG ACT CAG ACG GAA ACG CTT GAA CCG CGA GGT CGC TCC GGA CAC CAA TTT GAC TCG
(leu thr thr gln thr glu thr leu glu pro arg gly arg ser gly his gln phe asp ser
61/21   SEQ ID NO: 166   91/31
GCT CTT TGG CAA TTG AAG GTG AGC TGC GAG CAG CCG GGT GAC CGC ATC GTT GGC CTT GCC
ala leu trp gln leu lys val ser cys glu gln pro gly asp arg ile val gly leu ala
121/41   151/51
ATC AAT CGC CGG CTC GCG GAC GTA GAT AAT CAG CTC ACC GTT GGG ACC GAC CTC GAC CAG
ile asn arg arg leu ala asp val asp asn gln leu thr val gly thr asp leu asp gln
181/61   211/71
GGG TCC TTT GTG ACT GCC GGG CTT GAC GCG GAC GAC CAC AGA GTC GGT CAT CGC CTA AGG
gly ser phe val thr ala gly leu asp ala asp asp his arg val gly his arg leu arg
241/81   271/91
CTA CCG TTC TGA CCT GGG GCT GCG TGG GCG CCG ACG ACG TGA GGC ACG TCA TGT CTC AGC
leu pro phe)OPA(pro gly ala ala trp ala pro thr thr)OPA(gly thr ser cys leu ser
301/101   SEQ ID NO: 167   331/111   SEQ ID NO: 168
GGC CCA CCG CCA CCT CGG TCG CCG GCA GTA TGT CAG CAT GTG CAG ATG ACT CCA CGC AGC
gly pro pro pro pro arg ser pro ala val cys gln his val gln met thr pro arg ser
361/121   391/131
CTT GTT CGC ATC GTT GGT GTC GTG GTT GCG ACG ACC TTG GCG CTG GTG AGC GCA CCC GCC
leu val arg ile val gly val val val ala thr thr leu ala leu val ser ala pro ala
421/141
GGC GGT CGT GCC GCG CAT GCG GAT C
gly gly arg ala ala his ala asp)

SEQ ID NOS:165-168

FIG. 10A

SEQ ID NO: 169   32/11
TAA CGA CTC AGA CGG AAA CGC TTG AAC CGC GAG GTC GCT CCG GAC ACC AAT TTG ACT CGG
OCH(arg leu arg arg lys arg leu asn arg glu val ala pro asp thr asn leu thr arg
62/21   SEQ ID NO: 170   92/31
CTC TTT GGC AAT TGA AGG TGA GCT GCG AGC AGC CGG GTG ACC GCA TCG TTG GCC TTG CCA
leu phe gly asn)OPA arg OPA(ala ala ser ser arg val thr ala ser leu ala leu pro
122/41   SEQ ID NO: 171   152/51
TCA ATC GCC GGC TCG CGG ACG TAG ATA ATC AGC TCA CCG TTG GGA CCG ACC TCG ACC AGG
ser ile ala gly ser arg thr)AMB(ile ile ser ser pro leu gly pro thr ser thr arg
182/61   SEQ ID NO: 172   212/71
GGT CCT TTG TGA CTG CCG GGC TTG ACG CGG ACG ACC ACA GAG TCG GTC ATC GCC TAA GGC
gly pro leu)OPA(leu pro gly leu thr arg thr thr thr glu ser val ile ala)OCH(gly
242/81   SEQ ID NO: 173   272/91   SEQ ID NO: 174
TAC CGT TCT GAC CTG GGG CTG CGT GGG CGC CGA CGA CGT GAG GCA CGT CAT GTC TCA GCG
tyr arg ser asp leu gly leu arg gly arg arg arg arg glu ala arg his val ser ala
302/101   332/111
GCC CAC CGC GAC CTC GGT CGC CGG CAG TAT GTC AGC ATG TGC AGA TGA CTC CAC GCA GCC
ala his arg his leu gly arg arg gln tyr val ser met cys arg)OPA(leu his ala ala
362/121   392/131 SEQ ID NO: 175
TTG TTC GCA TCG TTG GTG TCG TGG TTG CGA CGA CCT TGG CGC TGG TGA GCG CAC CCG CCG
leu phe ala ser leu val ser trp leu arg arg pro trp arg trp)OPA(ala his pro pro
422/141   SEQ ID NO: 176
GCG GTC GTG CCG CGC ATG CGG ATC
ala val val pro arg met arg Ile)

SEQ ID NOS:170-176

FIG. 10B

SEQ ID NO: 177   SEQ ID NO: 179   33/11
AAC GAC TCA GAC GGA AAC GCT TGA ACC GCG AGG TCG CTC CGG ACA CCA ATT TGA CTC GGC
(asn asp ser asp gly asn ala)OPA(thr ala arg ser leu arg thr pro ile)OPA(leu gly
63/21   SEQ ID NO: 178   93/31   SEQ ID NO: 180
TCT TTG GCA ATT GAA GGT GAG CTG CGA GCA GCC GGG TGA CCG CAT CGT TGG CCT TGC CAT
ser leu ala ile glu gly glu leu arg ala ala gly)OPA(pro his arg trp pro cys his
123/41   153/51   SEQ ID NO: 181
CAA TCG CCG GCT CGC GGA CGT AGA TAA TCA GCT CAC CGT TGG ACG ACC TCG ACC AGG G
gln ser pro ala arg gly arg arg)OCH(ser ala his arg trp asp arg pro arg pro gly
183/61   SEQ ID NO: 182   213/71
GTC CTT TGT GAC TGC CGG GCT TGA CGC GGA CGA CCA CAG AGT CGG TCA TCG CCT AAG GCT
val leu cys asp cys arg ala)OPA(arg gly arg pro gln ser arg ser ser pro lys ala
243/81   SEQ ID NO: 183   273/91
ACC GTT CTG ACC TGG GGC TGC GTG GGC GCC GAC GAC GTG AGG CAC GTC ATG TCT CAG CGG
thr val leu thr trp gly cys val gly ala asp asp val arg his val met ser gln arg
303/101   333/111
CCC ACC GCC ACC TCG GTC GCC GGC AGT ATG TCA GCA TGT GCA GAT GAC TCC ACG CAG CCT
pro thr ala thr ser val ala gly ser met ser ala cys ala asp asp ser thr gln pro
363/121   393/131
TGT TCG CAT CGT TGG TGT CGT GGT TGC GAC GAC CTT GGC GCT GGT GAG CGC ACC CGC CGG
cys ser his arg trp cys arg gly cys asp asp leu gly ala gly glu arg thr arg arg
423/141
CGG TCG TGC CGC GCA TGC GGA TC
arg ser cys arg ala cys gly)

SEQ ID NOS:177-183

FIG. 10C

SEQ ID NO: 184   31/11
```
CCC GAA GAG GTC CCC CGT TTT GTT AAT TTT TAA AAA ATT TGT GTC ACA AAC CGG GGT ACC
(pro glu glu val pro arg phe val asp phe)OCH(lys ile cys val thr lys arg gly thr
61/21      SEQ ID NO:  185                91/31    SEQ ID NO:  186
AAG GCA TAA AAC CTA GTA CCT GGG GCG GCG GAT TCA ACG AAA ACC GAG TGG GGG TAG TCA
lys ala)OCH(asn leu val pro gly ala ala asp ser thr lys thr glu trp gly)AMB(ser
121/41         SEQ ID NO:  187           151/51                 SEQ ID NO:  188
GGG GCG TGC ATT CCG ACG ACC CTG TAC GAC CCG CTG GTG GCA ACG CCG ATG AGT GCG CCG
gly ala cys ile pro thr thr leu tyr asp pro leu val ala thr pro met ser ala pro
181/61                                   211/71
ACG AAG GCC GAG CGA CGG GCT GCC GGC GCT GAC CGC CGC GGA AGC CGC CGA GTG CAT GGT
thr lys ala glu arg arg ala ala gly ala asp arg arg gly ser arg arg val asp gly
241/81                                   271/91
CAC CAC CGC CCG CAC CCG ACC GGT ACG GAT CGC GCC TCG GGT TAC CGT CGC CGT CAA CGC
his his arg pro his pro thr gly thr asp arg ala ser gly tyr arg arg arg gln arg
301/101                                  331/111
GCT GGA CAG CAT CGG TCC CCG CTG GGT CAA TGC ACT CAT GCA GCG CCG CAA CGA ACA GCT
ala gly gln his arg ser pro leu gly gln cys thr his ala ala pro gln arg thr ala
361/121                                  391/131
CAA CCC TTG AAC CGG GTC CCG GCC TGC CGA CCC TCG GCC GCC GGC GTG CCG CTA CGT GAT
gln pro leu asn arg val pro ala cys arg pro ser ala ala gly val pro leu arg asp
421/141                                  451/151
AGA CAC AGG GCC ATG GAA ATC CTG GCC AGC CGG ATG CTA CTT CGG CCG GCG GAC TAT CAG
arg his arg ala met glu ile leu ala ser arg met leu leu arg pro ala asp tyr gln
481/161
CGG TCG CTG AGC TTC TAC CGT GAC CAG ATC
arg ser leu ser phe tyr arg asp gln ile)
```

SEQ ID NOS:184-188

FIG. 11A

SEQ ID NO: 189   32/11
```
CCG AAG AGG TCC CCC GTT TTG TTA ATT TTT AAA AAA TTT GTG TCA CAA AGC GGG GTA CCA
(pro lys arg ser pro val leu leu ile phe lys lys phe val ser gln ser gly val pro
62/21    SEQ ID NO:  190                 92/31
AGG CAT AAA ACC TAG TAC CTG GGG CGG CGG ATT CAA CGA AAA CCG AGT GGG GGT AGT CAG
arg his lys thr)AMB(tyr leu gly arg arg ile gln arg lys pro ser gly gly ser gln
122/41          SEQ ID NO:  191          152/51
GGG CGT GCA TTC CGA CGA CCC TGT ACG ACC CGC TGG TGG CAA CGC CGA TGA GTG CGC CGA
gly arg ala phe arg arg pro cys thr thr arg trp trp gln arg arg)OPA(val arg arg
182/61                                   212/71        SEQ ID NO:  192
CGA AGG CCG AGC GAC GGG CTG CCG GCG CTG ACC GCC GCG GAA GCC GCC GAG TGG ATG GTC
arg arg pro ser asp gly leu pro ala leu thr ala ala glu ala ala glu trp met val
242/81                                   272/91
ACC ACC GCC CGC ACC CGA CCG GTA CGG ATC GCG CCT CGG GTT ACC GTC GCC GTC ACC GCG
thr thr ala arg thr arg pro val arg ile ala pro arg val thr val ala val asn ala
302/101                                  332/111
CTG GAC AGC ATC GGT CCC CGC TGG GTC AAT GCA CTC ATG CAG CGC CGC AAC GAA CAG CTC
leu asp ser ile gly pro arg trp val asn ala leu met gln arg arg asn glu gln leu
362/121                                  392/131
AAC CCT TGA ACC GGG TCC CGG CCT GCC GAC CCT CGG CCG CCG GCG TGC CGC TAC GTG ATA
asn pro)OPA(thr gly ser arg pro ala asp pro arg pro pro ala cys arg tyr val ile
422/141        SEQ ID NO:  193           452/151
GAC ACA GGG CCA TGG AAA TCC TGG CCA GCC GGA TGC TAC TTC GCC GGC GGA CTA TCA GC
asp thr gly pro trp lys ser trp pro ala gly cys tyr phe gly arg arg thr ile ser
482/161
GGT CGC TGA GCT TCT ACC GTG ACC AGA TC
gly ar) OPA(ala ser thr val thr arg)
                SEQ ID NO:  194
```

SEQ ID NOS:189-194

FIG. 11B

SEQ ID NO: 195   33/11
CGA AGA GGT CCC CCG TTT TGT TAA TTT TTA AAA AAT TTG TGT CAC AAA GCG GGG TAC CAA
(arg arg gly pro pro phe cys)OCH(phe leu lys asn leu cys his lys ala gly tyr gln
63/21 SEQ ID NO: 196   93/31 SEQ ID NO: 197
GGC ATA AAA CCT AGT ACC TGG GGC GGC GGA TTC AAC GAA AAC CGA GTG GGG GTA GTC AGG
gly ile lys pro ser thr trp gly gly gly phe asn glu asn arg val gly val val arg
123/41   153/51
GGC GTG CAT TCC GAC GAC CCT GTA CGA CCC GCT GGT GGC AAC GCC GAT GAG TGC GCC GAC
gly val his ser asp asp pro val arg pro ala gly gly asn ala asp glu cys ala asp
183/61   213/71
GAA GGC CGA GCG ACG GGC TGC CGG CGC TGA CCG CCG CGG AAG CCG CCG AGT GGA TGG TCA
glu gly arg ala thr gly cys arg arg)OPA(pro pro arg lys pro pro ser gly trp ser
243/81   SEQ ID NO: 198 273/91
CCA CCG CCC GCA CCC GAC CGG TAC GGA TCG CGC CTC GGG TTA CCG TCG CCG TCA ACG CGC
pro pro pro ala pro asp arg tyr gly ser arg leu gly leu pro ser pro ser thr arg
303/101   333/111
TGG ACA GCA TCG GTC CCC GCT GGG TCA ATG CAC TCA TGC AGC GCC GCA ACG AAC AGC TCA
trp thr ala ser val pro ala gly ser met his ser cys ser ala ala thr asn ser ser
363/121   393/131
ACC CTT GAA CCG GGT CCC GGC CTG CCG ACC CTC GGC CGC CGG CGT GCC GCT ACG TGA TAG
thr leu glu pro gly pro gly leu pro thr leu gly arg arg arg ala ala thr)OPA AMB
423/141   453/151
ACA CAG GGC CAT GGA AAT CCT GGC CAG CCG GAT GCT ACT TCG GCC GGC GGA CTA TCA GCG
(thr gln gly his gly asn pro gly gln pro asp ala thr ser ala gly gly leu ser ala
483/161 SEQ ID NO: 199
GTC GCT GAG CTT CTA CCG TGA CCA GAT C
val ala glu leu leu pro)OPA pro asp

SEQ ID NOS:195-199

FIG. 11C

part of the nucleotide sequence of Seq11
1/1 SEQ ID NO: 200   31/11
CGT CGC CGT CAA CGC GCT GGA CAG CAT CGG TCC CCG CTG GGT CAA TGC ACT CAT GCA GCG
(arg arg arg gln arg ala gly gln his arg ser pro leu gly gln cys thr his ala ala
61/21 SEQ ID NO: 201   91/31
CCG CAA CGA ACA GCT CAA CCC TTG AAC CGG GTC CCG GCC TGC CGA CCC TCG GCC GCC GGC
pro gln arg thr ala gln pro leu asn arg val pro ala cys arg pro ser ala ala gly
121/41   151/51
GTG CCG CTA CGT GAT AGA CAC AGG GCC ATG GAA ATC CTG GCC AGC CGG ATG CTA CTT CGG
val pro leu arg asp arg his arg ala met glu ile leu ala ser arg met leu leu arg
181/61   211/71
CCG GCG GAC TAT CAG CGG TCG CTG AGC TTC TAC CGT GAC CAG ATC
pro ala asp tyr gln arg ser leu ser phe tyr arg asp gln ile)

SEQ ID NOS:200-201

FIG. 11A'

```
1/1    SEQ ID NO:  202                  31/11
GTC GCC GTC AAC GCG CTG GAC AGC ATC GGT CCC CGC TGG GTC AAT GCA CTC ATG CAG CGC
(val ala val asn ala leu asp ser ile gly pro arg trp val asn ala leu met gln arg
61/21    SEQ ID NO:  203                  91/31
CGC AAC GAA CAG CTC AAC CCT TGA ACC GGG TCC CGG CCT GCC GAC CCT CGG CCG CCG GCG
arg asn glu gln leu asn pro)OPA(thr gly ser arg pro ala asp pro arg pro pro ala
121/41       SEQ ID NO:  204           151/51
TGC CGC TAC GTG ATA GAC ACA GGG CCA TGG AAA TCC TGG CCA GCC GGA TGC TAC TTC GGC
cys arg tyr val ile asp thr gly pro trp lys ser trp pro ala gly cys tyr phe gly
181/61                                 211/71
CGG CGG ACT ATC AGC GGT CGC TGA GCT TCT ACC GTG ACC AGA TC
arg arg thr ile ser gly arg)OPA(ala ser thr val thr arg)
          SEQ ID NO:  205
```

SEQ ID NOS:202-205

FIG. 11B'

```
1/1    SEQ ID NO:  206                  31/11
TCG CCG TCA ACG CGC TGG ACA GCA TCG GTC CCC GCT GGG TCA ATG CAC TCA TGC AGC GCC
(ser pro ser thr arg trp thr ala ser val pro ala gly ser met his ser cys ser ala
61/21    SEQ ID NO:  207                  91/31
GCA ACG AAC AGC TCA ACC CTT GAA CCG GGT CCC GGC CTG CCG ACC CTC GGC CGC CGG CGT
ala thr asn ser ser thr leu glu pro gly pro gly leu pro thr leu gly arg arg arg
121/41                                 151/51
GCC GCT ACG TGA TAG ACA CAG GGC CAT GGA AAT CCT GGC CAG CCG GAT GCT ACT TCG GCC
ala ala thr)OPA AMB(thr gln gly his gly asn pro gly gln pro asp ala thr ser ala
181/61             SEQ ID NO:  208    211/71
GGC GGA CTA TCA GCG GTC GCT GAG CTT CTA CCG TGA CCA GAT C
gly gly leu ser ala val ala glu leu leu pro)OPA pro asp
```

SEQ ID NOS:206-208

FIG. 11C'

sequence Rv0546c predicted by Cole et al. (Nature 393:537-544) and containing
Seq11A'

```
1/1    SEQ ID NO:  209                  31/11
atg gaa atc ctg gcc agc cgg atg cta ctt cgg ccg gcg gac tat cag cgg tcg ctg agc
(Met glu ile leu ala ser arg met leu leu arg pro ala asp tyr gln arg ser leu ser
61/21    SEQ ID NO:  210                  91/31
ttc tac cgt gac cag atc ggg ctg gcg att gcc cgt gaa tac ggg gcc ggc aca gtg ttt
phe tyr arg asp gln ile gly leu ala ile ala arg glu tyr gly ala gly thr val phe
121/41                                 151/51
ttc gcc ggt cag tca ctg ctc gaa ctg gcc ggt tac ggc gag ccg gac cat tcg cgg gga
phe ala gly gln ser leu leu glu leu ala gly tyr gly glu pro asp his ser arg gly
181/61                                 211/71
cct ttt ccc ggc gcg ctg tgg ctg cag gtg cgc gac ctc gag gct acc cag acc gag ctg
pro phe pro gly ala leu trp leu gln val arg asp leu glu ala thr gln thr glu leu
241/81                                 271/91
gtc agc cga ggc gtg tcg atc gct cgc gag ccc cgc cgc gaa ccg tgg ggc ctg cac gag
val ser arg gly val ser ile ala arg glu pro arg arg glu pro trp gly leu his glu
301/101                                331/111
atg cat gtg acc gac cca gac ggg atc aca ctg ata ttc gtc gag gtt ccc gag ggt cac
met his val thr asp pro asp gly ile thr leu ile phe val glu val pro glu gly his
361/121
ccg ctg cgt aca gac acc cgg gcg tga
pro leu arg thr asp thr arg ala)OPA
```

SEQ ID NOS:209-210
FIG. 11D

ORF predicted by Cole et al. (Nature 393:537-544) and containing Rv0546c

1/1 ⟋SEQ ID NO: 211   31/11
tag tca ggg cgt gca ttc gac gac gct gta cta ccc gct ggt ggc aac tcc gat gat tgc
AMB(ser gly arg ala phe asp asp ala val leu pro ala gly gly asn ser asp asp cys
61/21 ⟍SEQ ID NO: 212   91/31
gcc gac gaa ggc cta cga cgg gct gcc ggc gct gac cgc cgc gga agc cgc cga gtg gat
ala asp glu gly leu arg arg ala ala gly ala asp arg arg gly ser arg arg val asp
121/41   151/51
ggt cac cgc cgc ccg cac ccg acc ggt gcg gat cgc gcc tcg ggt tgc cgt cgc cgt caa
gly his arg arg pro his pro thr gly ala asp arg ala ser gly cys arg arg arg gln
181/61   211/71
cgc gct gga cag cat cgg tcc ccg ctg ggt caa tgc act cat gca gcg ccg caa cga aca
arg ala gly gln his arg ser pro leu gly gln cys thr his ala ala pro gln arg thr
241/81   271/91
gct caa ccc ttg aac cgg gtc ccg gcc tgc cga ccc tcg gcc gcc ggc gtg ccg cta cgt
ala gln pro leu asn arg val pro ala cys arg pro ser ala ala gly val pro leu arg
301/101   331/111
gat aga cac agg gcc atg gaa atc ctg gcc agc cgg atg cta ctt cgg ccg gcg gac tat
asp arg his arg ala met glu ile leu ala ser arg met leu leu arg pro ala asp tyr
361/121   391/131
cag cgg tcg ctg agc ttc tac cgt gac cag atc ggg ctg gcg att gcc cgt gaa tac ggg
gln arg ser leu ser phe tyr arg asp gln ile gly leu ala ile ala arg glu tyr gly
421/141   451/151
gcc ggc aca gtg ttt ttc gcc ggt cag tca ctg ctc gaa ctg gcc ggt tac ggc gag ccg
ala gly thr val phe phe ala gly gln ser leu leu glu leu ala gly tyr gly glu pro
481/161   511/171
gac cat tcg cgg gga cct ttt ccc ggc gcg ctg tgg ctg cag gtg cgc gac ctc gag gct
asp his ser arg gly pro phe pro gly ala leu trp leu gln val arg asp leu glu ala
541/181   571/191
acc cag acc gag ctg gtc agc cga ggc gtg tcg atc gct cgc gag ccc cgc cgc gaa ccg
thr gln thr glu leu val ser arg gly val ser ile ala arg glu pro arg arg glu pro
601/201   631/211
tgg ggc ctg cac gag atg cat gtg acc gac cca gac ggg atc aca ctg ata ttc gtc gag
trp gly leu his glu met his val thr asp pro asp gly ile thr leu ile phe val glu
661/221   691/231
gtt ccc gag ggt cac ccg ctg cgt aca gac acc cgg gcg tga
val pro glu gly his pro leu arg thr asp thr arg ala)OPA

SEQ ID NOS:211-212

FIG. 11E

1/1 — SEQ ID NO: 213   31/11
gac cga agg gat ttc gcg act aac tcg gcc tgt aag gca acg cga ggt ctt cat gcc gag
(asp arg arg asp phe ala thr asn ser ala cys lys ala thr arg gly leu his ala glu
61/21 — SEQ ID NO: 214   91/31
gac gta gac agg aag aga cag gga agc tga tga cgt cgc gta ccg gac cgc cat tct gtc
asp val asp arg lys arg gln gly ser)OPA OPA(arg arg val pro asp arg his ser val
121/41   151/51 — SEQ ID NO: 215
gag tct ttc cga gtt cag caa caa tcg aca cag aag cgg gga cca gac cgg gag gac gac
glu ser phe arg val gln gln gln ser thr gln lys arg gly pro asp arg glu asp asp
181/61   211/71
gcg gcc cgg gcc gct tcg ggc cga gtg tct gag taa gac cag agt cac ggg tcc gtg tgt
ala ala arg ala ala ser gly arg val ser glu)OCH(asp gln ser his gly ser val cys
241/81   271/91   — SEQ ID NO: 216
gac aac cgc gcg gaa ttc aat cgg atg gcg ggc ggg acc gga ttg cgc cgg tca ccg agg
asp asn arg ala glu phe asn arg met ala gly gly thr gly leu arg arg ser pro arg
301/101
aac ctc cgg agt gat c
asn leu arg ser asp)

SEQ ID NOS:213-216

FIG. 12A

1/1 ____ SEQ ID NO: 217   31/11
acc gaa ggg att tcg cga cta act cgg cct gta agg caa cgc gag gtc ttc atg ccg agg
(thr glu gly ile ser arg leu thr arg pro val arg gln arg glu val phe met pro arg
61/21 — SEQ ID NO: 218   91/31
acg tag aca gga aga gac agg gaa gct gat gac gtc gcg tac cgg acc gcc att ctg tcg
thr)AMB(thr gly arg asp arg glu ala asp asp val ala tyr arg thr ala ile leu ser
121/41   — SEQ ID NO: 219   151/51
agt ctt tcc gag ttc agc aac aat cga cac aga agc ggg gac cag acc ggg agg acg acg
ser leu ser glu phe ser asn asn arg his arg ser gly asp gln thr gly arg thr thr
181/61   211/71
cgg ccc ggg ccg ctt cgg gcc gag tgt ctg agt aag acc aga gtc acg ggt ccg tgt gtg
arg pro gly pro leu arg ala glu cys leu ser lys thr arg val thr gly pro cys val
241/81   271/91
aca acc gcg cgg aat tca atc gga tgg cgg gcg gga ccg gat tgc gcc ggt cac cga gga
thr thr ala arg asn ser ile gly trp arg ala gly pro asp cys ala gly his arg gly
301/101
acc tcc gga gtg atc
thr ser gly val ile)

SEQ ID NOS:217-219

FIG. 12B

1/1 SEQ ID NO: 220   31/11
ccg aag gga ttt cgc gac taa ctc ggc ctg taa ggc aac gcg agg tct tca tgc cga gga
(pro lys gly phe arg asp)OCH leu gly leu OCH(gly asn ala arg ser ser cys arg gly
61/21 SEQ ID NO: 221   91/31 SEQ ID NO: 222
cgt aga cag gaa gag aca ggg aag ctg atg acg tcg cgt acc gga ccg cca ttc tgt cga
arg arg gln glu glu thr gly lys leu met thr ser arg thr gly pro pro phe cys arg
121/41   151/51
gtc ttt ccg agt tca gca aca atc gac aca gaa gcg ggg acc aga ccg gga gga cga cgc
val phe pro ser ser ala thr ile asp thr glu ala gly thr arg pro gly gly arg arg
181/61   211/71
ggc ccg ggc cgc ttc ggg ccg agt gtc tga gta aga cca gag tca cgg gtc cgt gtg tga
gly pro gly arg phe gly pro ser val)OPA(val arg pro glu ser arg val arg val)OPA
241/81   SEQ ID NO: 223   271/91
caa ccg cgc gga att caa tcg gat ggc ggg cgg gac cgg att gcg ccg gtc acc gag gaa
(gln pro arg gly ile gln ser asp gly gly arg asp arg ile ala pro val thr glu glu
301/101 SEQ ID NO: 224
cct ccg gag tga tc
pro pro glu)OPA

SEQ ID NOS:220-224

FIG. 12C

1/1 SEQ ID NO: 225   31/11
GGG ATT TCG TTG CCC GAT GGA TTG TTT GTA CGG TTT GGG AAA AAC ACT TGA AGT CCT TTT
(gly ile ser leu pro asp gly leu phe val arg phe gly lys asn thr)OPA(ser pro phe
61/21 SEQ ID NO: 226   91/31 SEQ ID NO: 227
TAT TGG CAA TGC TGG AAA TGG ACA TTC CAA TAT TGC GCG AAT TAA CCG AAC ACG GTG AGG
tyr trp gln cys trp lys trp thr phe gln tyr cys ala asn)OCH(pro asn thr val arg
121/41   151/51 SEQ ID NO: 923
GGG GGG CAA GCG TTT GTA CCG GGG CCA GCA AGC GCC GCC GAC CGG TTG ACC GAA GCC AGC
gly gly gln ala phe val pro gly pro ala ser ala ala asp arg leu thr glu ala ser
181/61   211/71
ATG TTG TTG TGT CAG CGC GGG CTT GGT CTC GAT GTC CCG GCC TTG GCT GGA CCC GCT TCT
met leu leu cys gln arg gly leu gly leu asp val pro ala leu ala gly pro ala ser
241/81   271/91
TCA AAA CAG GTT GAA CTT AAC GAC TCA AGA ACG GAA ACG CTT GAA CCG CGA CGT CGC TCC
ser lys gln val glu leu asn asp ser arg thr glu thr leu glu pro arg arg arg ser
301/101   331/111
GGA CAC CAA TTT GAC TCG GCT CTT TGG CAA TTG AAG GTG AGC TGC GAG CAG CCG GGT GAC
gly his gln phe asp ser ala leu trp gln leu lys val ser cys glu gln pro gly asp
361/121   391/131
CGC ATC GTT GGC CTT GCC ATC AAT CGC CGG CTC GCG GAC GTA GAT AAT CAG CTC ACC GTT
arg ile val gly leu ala ile asn arg arg leu ala asp val asp asn gln leu thr val
421/141   451/151
GGG ACC GAC CTC GAC CAG GGG TCC TTT GTG ACT GCC GGG CTT GAC GCG GAC GAC CAC AGA
gly thr asp leu asp gln gly ser phe val thr ala gly leu asp ala asp asp his arg
481/161   511/171
GTC GGT CAT CGC CTA AGG CTA CCG TTC TGA CCT GGG GCT GCG TGG GCG CCG ACG ACG TGA
val gly his arg leu arg leu pro phe)OPA(pro gly ala ala trp ala pro thr thr)OPA
541/181   SEQ ID NO: 924   571/191
GGC ACG TCA TGT CTC AGC GGC CCA CCG CCA CCT CGG TCG CCG GCA GTA TGT CAG CAT GTG
(gly thr ser cys leu ser gly pro pro pro arg ser pro ala val cys gln his val
601/201 SEQ ID NO: 925   631/211
CAG ATG ACT CCA CGC AGC CTT GTT CGC ATC GTT GGT GTC GTG GTT GCG ACG ACC TTG GCG
gln met thr pro arg ser leu val arg ile val gly val val val ala thr thr leu ala
661/221   691/231
CTG GTG AGC GCA CCC GCC GGC GGT CGT GCC GCG CAT GCG GAT C
leu val ser ala pro ala gly gly arg ala ala his ala asp)

SEQ ID NOS:220-227,923-925

FIG. 13A

SEQ ID NO: 228   32/11
GGA TTT CGT TGC CCG ATG GAT TGT TTG TAC GGT TTG GGA AAA ACA CTT GAA GTC CTT TTT
(gly phe arg cys pro met asp cys leu tyr gly leu gly lys thr leu glu val leu phe
62/21   SEQ ID NO: 229   92/31
ATT GGC AAT GCT GGA AAT GGA CAT TCC AAT ATT GCG CGA ATT AAC CGA ACA CGG TGA GGG
ile gly asn ala gly asn gly his ser asn ile ala arg ile asn arg thr arg)OPA(gly
122/41   152/51   SEQ ID NO: 230
GGG GGC AAG CGT TTG TAC CGG GGC CAG CAA GCG CCG CCG ACC GGT TGA CCG AAG CCA GCA
gly gly lys arg leu tyr arg gly gln gln ala pro pro thr gly)OPA(pro lys pro ala
182/61   212/71   SEQ ID NO: 231
TGT TGT TGT GTC AGC GCG GGC TTG GTC TCG ATG TCC CGG CCT TGG CTG GAC CCG CTT CTT
cys cys cys val ser ala gly leu val ser met ser arg pro trp leu asp pro leu leu
242/81   272/91
CAA AAC AGG TTG AAC TTA ACG ACT CAA GAA CGG AAA CGC TTG AAC CGC GAC GTC GCT CCG
gln asn arg leu asn leu thr thr gln glu arg lys arg leu asn arg asp val ala pro
302/101   332/111
GAC ACC AAT TTG ACT CGG CTC TTT GGC AAT TGA AGG TGA GCT GCG AGC AGC CGG GTG ACC
asp thr asn leu thr arg leu phe gly asn)OPA arg OPA(ala ala ser ser arg val thr
362/121   392/131   SEQ ID NO: 232
GCA TCG TTG GCC TTG CCA TCA ATC GCC GGC TCG CGG ACG TAG ATA ATC AGC TCA CCG TTG
ala ser leu ala leu pro ser ile ala gly ser arg thr)AMB(ile ile ser ser pro leu
422/141   452/151   SEQ ID NO: 233
GGA CCG ACC TCG ACC AGG GGT CCT TTG TGA CTG CCG GGC TTG ACG CGG ACG ACC ACA GAG
gly pro thr ser thr arg gly pro leu)OPA(leu pro gly leu thr arg thr thr thr glu
482/161   SEQ ID NO: 234 512/171
TCG GTC ATC GCC TAA GGC TAC CGT TCT GAC CTG GGG CTG CGT GGG CGC CGA CGA CGT GAG
ser val ile ala)OCH(gly tyr arg ser asp leu gly leu arg gly arg arg arg arg glu
542/181   SEQ ID NO: 235 572/191
GCA CGT CAT GTC TCA GCG GCC CAC CGC CAC CTC GGT CGC CGG CAG TAT GTC AGC ATG TGC
ala arg his val ser ala ala his arg his leu gly arg arg gln tyr val ser met cys
602/201   632/211
AGA TGA CTC CAC GCA GCC TTG TTC GCA TCG TTG GTG TCG TGG TTG CGA CGA CCT TGG CGC
arg)OPA(leu his ala ala leu phe ala ser leu val ser trp leu arg arg pro trp arg
662/221   SEQ ID NO: 236   692/231
TGG TGA GCG CAC CCG CCG GCG GTC GTG CCG CGC ATG CGG ATC
trp)OPA(ala his pro pro ala val val pro arg met arg ile)
           SEQ ID NO: 237

SEQ ID NOS:228-237

FIG. 13B

SEQ ID NO: 238   33/11
GAT TTC GTT GCC CGA TGG ATT GTT TGT ACG GTT TGG GAA AAA CAC TTG AAG TCC TTT TTA
(asp phe val ala arg trp ile val cys thr val trp glu lys his leu lys ser phe leu
63/21   SEQ ID NO: 239   93/31
TTG GCA ATG CTG GAA ATG GAC ATT CCA ATA TTG CGC GAA TTA ACC GAA CAC GGT GAG GGG
leu ala met leu glu met asp ile pro ile leu arg glu leu thr glu his gly glu gly
123/41   153/51
GGG GCA AGC GTT TGT ACC GGG GCC AGC AAG CGC CGC CGA CCG GTT GAC CGA AGC CAG CAT
gly ala ser val cys thr gly ala ser lys arg arg arg pro val asp arg ser gln his
183/61   213/71
GTT GTT GTG TCA GCG CGG GCT TGG TCT CGA TGT CCC GGC CTT GGC TGG ACC CGC TTC TTC
val val val ser ala arg ala trp ser arg cys pro gly leu gly trp thr arg phe phe
243/81   273/91
AAA ACA GGT TGA ACT TAA CGA CTC AAG AAC GGA AAC GCT TGA ACC GCG ACG TCG CTC CGG
lys thr gly)OPA thr OCH(arg leu lys asn gly asn ala)OPA(thr ala thr ser leu arg
303/101 SEQ ID NO: 240   333/111   SEQ ID NO: 241
ACA CCA ATT TGA CTC GGC TCT TTG GCA ATT GAA GGT GAG CTG CGA GCA GCC GGG TGA CCG
thr pro ile)OPA(leu gly ser leu ala ile glu gly glu leu arg ala ala gly)OPA(pro
363/121   SEQ ID NO: 242   393/131   SEQ ID NO: 243
CAT CGT TGG CCT TGC CAT CAA TCG CCG GCT CGC GGA CGT AGA TAA TCA GCT CAC CGT TGG
his arg trp pro cys his gln ser pro ala arg gly arg arg)OCH(ser ala his arg trp
423/141   453/151   SEQ ID NO: 244
GAC CGA CCT CGA CCA GGG GTC CTT TGT GAC TGC CGG GCT TGA CGC GGA CGA CCA CAG AGT
asp arg pro arg pro gly val leu cys asp cys arg ala)OPA(arg gly arg pro gln ser
483/161   513/171   SEQ ID NO: 245
CGG TCA TCG CCT AAG GCT ACC GTT CTG ACC TGG GGC TGC GTG GGC GCC GAC GAC GTG AGG
arg ser ser pro lys ala thr val leu thr trp gly cys val gly ala asp asp val arg
543/181   573/191
CAC GTC ATG TCT CAG CGG CCC ACC GCC ACC TCG GTC GCC GGC AGT ATG TCA GCA TGT GCA
his val met ser gln arg pro thr ala thr ser val ala gly ser met ser ala cys ala
603/201   633/211
GAT GAC TCC ACG CAG CCT TGT TCG CAT CGT TGG TGT CGT GGT TGC GAC GAC CTT GGC GCT
asp asp ser thr gln pro cys ser his arg trp cys arg gly cys asp asp leu gly ala
663/221   693/231
GGT GAG CGC ACC CGC CGG CGG TCG TGC CGC GCA TGC GGA TC
gly glu arg thr arg arg arg ser cys arg ala cys gly)

SEQ ID NOS:238-245

FIG. 13C

part of the nucleotide sequence of seq13A

```
1/1    SEQ ID NO:  246                    31/11
GGG TCC TTT GTG ACT GCC GGG CTT GAC GCG GAC GAC CAC AGA GTC GGT CAT CGC CTA AGG
(gly ser phe val thr ala gly leu asp ala asp asp his arg val gly his arg leu arg
61/21    SEQ ID NO:  247                    91/31
CTA CCG TTC TGA CCT GGG GCT GCG TGG GCG CCG ACG ACG TGA GGC ACG TCA TGT CTC AGC
leu pro phe)OPA(pro gly ala ala trp ala pro thr thr)OPA(gly thr ser cys leu ser
121/41          SEQ ID NO:  248          151/51          SEQ ID NO:  249
GGC CCA CCG CCA CCT CGG TCG CCG GCA GTA TGT CAG CAT GTG CAG ATG ACT CCA CGC AGC
gly pro pro pro pro arg ser pro ala val cys gln his val gln met thr pro arg ser
181/61                                   211/71
CTT GTT CGC ATC GTT GGT GTC GTG GTT GCG ACG ACC TTG GCG CTG GTG AGC GCA CCC GCC
leu val arg ile val gly val val val ala thr thr leu ala leu val ser ala pro ala
241/81
GGC GGT CGT GCC GCG CAT GCG GAT C
gly gly arg ala ala his ala asp)
```

SEQ ID NOS:246-249

FIG. 13A'

```
1/1   SEQ ID NO:  250                    31/11
GGT CCT TTG TGA CTG CCG GGC TTG ACG CGG ACG ACC ACA GAG TCG GTC ATC GCC TAA GGC
gly pro(leu OPA leu pro gly leu thr arg thr thr thr glu ser val ile ala)OCH(gly
61/21   SEQ ID NO:  251                    91/31          SEQ ID NO:  252
TAC CGT TCT GAC CTG GGG CTG CGT GGG CGC CGA CGA CGT GAG GCA CGT CAT GTC TCA GCG
tyr arg ser asp leu gly leu arg gly arg arg arg arg glu ala arg his val ser ala
121/41                                   151/51
GCC CAC CGC CAC CTC GGT CGC CGG CAG TAT GTC AGC ATG TGC AGA TGA CTC CAC GCA GCC
ala his arg his leu gly arg arg gln tyr val ser met cys arg)OPA(leu his ala ala
181/61                                   211/71          SEQ ID NO:  253
TTG TTC GCA TCG TTG GTG TCG TGG TTG CGA CGA CCT TGG CGC TGG TGA GCG CAC CCG CCG
leu phe ala ser leu val ser trp leu arg arg pro trp arg trp)OPA(ala his pro pro
241/81                                   SEQ ID NO:  254
GCG GTC GTG CCG CGC ATG CGG ATC
ala val val pro arg met arg ile)
```

SEQ ID NOS:250-254

FIG. 13B'

```
1/1   SEQ ID NO:  255                    31/11
GTC CTT TGT GAC TGC CGG GCT TGA CGC GGA CGA CCA CAG AGT CGG TCA TCG CCT AAG GCT
(val leu cys asp cys arg ala)OPA(arg gly arg pro gln ser arg ser ser pro lys ala
61/21    SEQ ID NO:  256                    91/31  SEQ ID NO:  257
ACC GTT CTG ACC TGG GGC TGC GTG GGC GCC GAC GAC GTG AGG CAC GTC ATG TCT CAG CGG
thr val leu thr trp gly cys val gly ala asp asp val arg his val met ser gln arg
121/41                                   151/51
CCC ACC GCC ACC TCG GTC GCC GGC AGT ATG TCA GCA TGT GCA GAT GAC TCC ACG CAG CCT
pro thr ala thr ser val ala gly ser met ser ala cys ala asp asp ser thr gln pro
181/61                                   211/71
TGT TCG CAT CGT TGG TGT CGT GGT TGC GAC GAC CTT GGC GCT GGT GAG CGC ACC CGC CGG
cys ser his arg trp cys arg gly cys asp asp leu gly ala gly glu arg thr arg arg
241/81
CGG TCG TGC CGC GCA TGC GGA TC
arg ser cys arg ala cys gly)
```

SEQ ID NOS:255-257

FIG. 13C'

sequence Rv1984c predicted by Cole et al. (Nature 393:537-544) and
containing seq13A'

```
1/1    SEQ ID NO:  258                    31/11
atg act cca cgc agc ctt gtt cgc atc gtt ggt gtc gtg gtt gcg acg acc ttg gcg ctg
(Met thr pro arg ser leu val arg ile val gly val val val ala thr thr leu ala leu
61/21    SEQ ID NO:  259                    91/31
gtg agc gca ccc gcc ggc ggt cgt gcc gcg cat gcg gat ccg tgt tcg gac atc gcg gtc
val ser ala pro ala gly gly arg ala ala his ala asp pro cys ser asp ile ala val
121/41                                      151/51
gtt ttc gct cgc ggc acg cat cag gct tct ggt ctt ggc gac gtc ggt gag gcg ttc gtc
val phe ala arg gly thr his gln ala ser gly leu gly asp val gly glu ala phe val
181/61                                      211/71
gac tcg ctt acc tcg caa gtt ggc ggg cgg tcg att ggg gtc tac gcg gtg aac tac cca
asp ser leu thr ser gln val gly gly arg ser ile gly val tyr ala val asn tyr pro
241/81                                      271/91
gca agc gac gac tac cgc gcg agc gcg tca aac ggt tcc gat gat gcg agc gcc cac atc
ala ser asp asp tyr arg ala ser ala ser asn gly ser asp asp ala ser ala his ile
301/101                                     331/111
cag cgc acc gtc gcc agc tgc ccg aac acc agg att gtg ctt ggt ggc tat tcg cag ggt
gln arg thr val ala ser cys pro asn thr arg ile val leu gly gly tyr ser gln gly
361/121                                     391/131
gcg acg gtc atc gat ttg tcc acc tcg gcg atg ccg ccc gcg gtg gca gat cat gtc gcc
ala thr val ile asp leu ser thr ser ala met pro pro ala val ala asp his val ala
421/141                                     451/151
gct gtc gcc ctt ttc ggc gag cca tcc agt ggt ttc tcc agc atg ttg tgg ggc ggc ggg
ala val ala leu phe gly glu pro ser ser gly phe ser ser met leu trp gly gly gly
481/161                                     511/171
tcg ttg ccg aca atc ggt ccg ctg tat agc tct aag acc ata aac ttg tgt gct ccc gac
ser leu pro thr ile gly pro leu tyr ser ser lys thr ile asn leu cys ala pro asp
541/181                                     571/191
gat cca ata tgc acc gga ggc ggc aat att atg gcg cat gtt tcg tat gtt cag tcg ggg
asp pro ile cys thr gly gly gly asn ile met ala his val ser tyr val gln ser gly
601/201                                     631/211
atg aca agc cag gcg gcg aca ttc gcg gcg aac agg ctc gat cac gcc gga tga
met thr ser gln ala ala thr phe ala ala asn arg leu asp his ala gly)OPA
```

SEQ ID NOS:258-259

FIG. 13D

Seq13F: ORF predicted by Cole et al. (Nature 393:537-544) and containing
Rv1984c

1/1 — SEQ ID NO: 260   31/11
tga ggc acg tca tgt ctc agc ggc cca ccg cca cct cgg tcg ccg gca gta tgt cag cat
OPA(gly thr ser cys leu ser gly pro pro pro pro arg ser pro ala val cys gln his
61/21 — SEQ ID NO: 261   91/31
gtg cag atg act cca cgc agc ctt gtt cgc atc gtt ggt gtc gtg gtt gcg acg acc ttg
val gln met thr pro arg ser leu val arg ile val gly val val val ala thr thr leu
121/41   151/51
gcg ctg gtg agc gca ccc gcc ggc ggt cgt gcc gcg cat gcg gat ccg tgt tcg gac atc
ala leu val ser ala pro ala gly gly arg ala ala his ala asp pro cys ser asp ile
181/61   211/71
gcg gtc gtt ttc gct cgc ggc acg cat cag gct tct ggt ctt ggc gac gtc ggt gag gcg
ala val val phe ala arg gly thr his gln ala ser gly leu gly asp val gly glu ala
241/81   271/91
ttc gtc gac tcg ctt acc tcg caa gtt ggc ggg cgg tcg att ggg gtc tac gcg gtg aac
phe val asp ser leu thr ser gln val gly gly arg ser ile gly val tyr ala val asn
301/101   331/111
tac cca gca agc gac gac tac cgc gcg agc gcg tca aac ggt tcc gat gat gcg agc gcc
tyr pro ala ser asp asp tyr arg ala ser ala ser asn gly ser asp asp ala ser ala
361/121   391/131
cac atc cag cgc acc gtc gcc agc tgc ccg aac acc agg att gtg ctt ggt ggc tat tcg
his ile gln arg thr val ala ser cys pro asn thr arg ile val leu gly gly tyr ser
421/141   451/151
cag ggt gcg acg gtc atc gat ttg tcc acc tcg gcg atg ccg ccc gcg gtg gca gat cat
gln gly ala thr val ile asp leu ser thr ser ala met pro pro ala val ala asp his
481/161   511/171
gtc gcc gct gtc gcc ctt ttc ggc gag cca tcc agt ggt ttc tcc agc atg ttg tgg ggc
val ala ala val ala leu phe gly glu pro ser ser gly phe ser ser met leu trp gly
541/181   571/191
ggc ggg tcg ttg ccg aca atc ggt ccg ctg tat agc tct aag acc ata aac ttg tgt gct
gly gly ser leu pro thr ile gly pro leu tyr ser ser lys thr ile asn leu cys ala
601/201   631/211
ccc gac gat cca ata tgc acc gga ggc ggc aat att atg gcg cat gtt tcg tat gtt cag
pro asp asp pro ile cys thr gly gly gly asn ile met ala his val ser tyr val gln
661/221   691/231
tcg ggg atg aca agc cag gcg gcg aca ttc gcg gcg aac agg ctc gat cac gcc gga tga
ser gly met thr ser gln ala ala thr phe ala ala asn arg leu asp his ala gly)OPA

SEQ ID NOS:260-261

FIG. 13E

SEQ ID NO: 262   31/11
CCA CCG GGG CTG GAG GGG CGA ATG TGC GCC GAA CGC CGT CGG CCA ACT TGG CCG CTG AGG
(pro pro gly leu glu gly arg met cys ala glu arg arg arg pro thr trp pro leu arg
61/21 SEQ ID NO: 263   91/31
GCG GCT GAT CCC CTG GCC CGA GAC GGG GCA AGC CAA TAG CGG CTC CAT CGG GCT TTG CTG
ala ala asp pro leu ala arg asp gly ala ser gln)AMB(arg leu his arg ala leu leu
121/41   151/51   SEQ ID NO: 264
GTA GCG GTT CGG CGG GAA CCG AGC GCC GAC GTT GTC GGT GCC CGG TGA TAT ATT GGG TCA
val ala val arg arg glu pro ser ala asp val val gly ala arg)OPA(tyr ile gly ser
181/61   211/71   SEQ ID NO: 265
GAC GGG TAT GGC GGC GAC TGA GGT GAT CTG CGA CAC GCC GCC GCG GTG CTC GAG CCA GGC
asp gly tyr gly gly asp)OPA(gly asp leu arg his ala ala ala val leu glu pro gly
241/81   SEQ ID NO: 266   271/91
TTA CGA CCA GGG AAT TTC GAA AAT GTT ATT CAG AAC ATC TTG TAT CTC TTC CTC CGT GCC
leu arg pro gly asn phe glu asn val ile gln asn ile leu tyr leu phe leu arg ala
301/101   331/111
ACC CCC TAG GTG TAG TGT TTT CGA GTA CCG GCA GAT CCC AGT TCA CCA GTC TCA CCA GAT
thr pro)AMB val AMB(cys phe arg val pro ala asp pro ser ser pro val ser pro asp)
                          SEQ ID NO: 267
C

SEQ ID NOS:262-267

FIG. 14A

SEQ ID NO: 268   32/11
CAC CGG GGC TGG AGG GGC GAA TGT GCG CCG AAC GCC GTC GGC CAA CTT GGC CGC TGA GGG
(his arg gly trp arg gly glu cys ala pro asn ala val gly gln leu gly arg)OPA(gly
62/21   SEQ ID NO: 269   92/31   SEQ ID NO: 270
CGG CTG ATC CCC TGG CCC GAG ACG GGG CAA GCC AAT AGC GGC TCC ATC GGG CTT TGC TGG
arg leu ile pro trp pro glu thr gly gln ala asn ser gly ser ile gly leu cys trp)
122/41   152/51
TAG CGG TTC GGC GGG AAC CGA GCG CCG ACG TTG TCG GTG CCC GGT GAT ATA TTG GGT CAG
AMB(arg phe gly gly asn arg ala pro thr leu ser val pro gly asp ile leu gly gln
182/61   SEQ ID NO: 271   212/71
ACG GGT ATG GCG GCG ACT GAG GTG ATC TGC GAC ACG CCG CCG CGG TGC TCG AGC CAG GCT
thr gly met ala ala thr glu val ile cys asp thr pro pro arg cys ser ser gln ala
242/81   272/91
TAC GAC CAG GGA ATT TCG AAA ATG TTA TTC AGA ACA TCT TGT ATC TCT TCC TCC GTG CCA
tyr asp gln gly ile ser lys met leu phe arg thr ser cys ile ser ser ser val pro
302/101   332/111
CCC CCT AGG TGT AGT GTT TTC GAG TAC CGG CAG ATC CCA GTT CAC CAG TCT CAC CAG ATC
pro pro arg cys ser val phe glu tyr arg gln ile pro val his gln ser his gln ile)

SEQ ID NOS:268-271

FIG. 14B

SEQ ID NO: 274   33/11

ACC GGG GCT GGA GGG GCG AAT GTG CGC CGA ACG CCG TCG GCC AAC TTG GCC GCT GAG GGC
(thr gly ala gly gly ala asn val arg arg thr pro ser ala asn leu ala ala glu gly
63/21 SEQ ID NO: 275   93/31

GGC TGA TCC CCT GGC CCG AGA CGG GGC AAG CCA ATA GCG GCT CCA TCG GGC TTT GCT GGT
gly)OPA(ser pro gly pro arg arg gly lys pro ile ala ala pro ser gly phe ala gly
123/41 SEQ ID NO: 276   153/51

AGC GGT TCG GCG GGA ACC GAG CGC CGA CGT TGT CGG TGC CCG GTG ATA TAT TGG GTC AGA
ser gly ser ala gly thr glu arg arg arg cys arg cys pro val ile tyr trp val arg
183/61   213/71

CGG GTA TGG CGG CGA CTG AGG TGA TCT GCG ACA CGC CGC CGC GGT GCT CGA GCC AGG CTT
arg val trp arg arg leu arg)OPA(ser ala thr arg arg arg gly ala arg ala arg leu
243/81   SEQ ID NO: 277 273/91

ACG ACC AGG GAA TTT CGA AAA TGT TAT TCA GAA CAT CTT GTA TCT CTT CCT CCG TGC CAC
thr thr arg glu phe arg lys cys tyr ser glu his leu val ser leu pro pro cys his
303/101   333/111

CCC CTA GGT GTA GTG TTT TCG AGT ACC GGC AGA TCC CAG TTC ACC AGT CTC ACC AGA TC
pro leu gly val val phe ser ser thr gly arg ser gln phe thr ser leu thr arg)

SEQ ID NOS:274-277

FIG. 14C

part of the nucleotide sequence of seq14A

1/1 SEQ ID NO: 278   31/11

TTT TCG AGT ACC GGC AGA TCC CAG GTT CAC CAG GTC TCA CCA GAT C
phe ser ser thr gly arg ser gln val his gln val ser pro asp
 SEQ ID NO: 279

SEQ ID NOS:278-279

FIG. 14A'

1/1 SEQ ID NO: 280   31/11

TGT TTT CGA GTA CCG GCA GAT CCC AGG TTC ACC AGG TCT CAC CAG ATC
cys phe arg val pro ala asp pro arg phe thr arg ser his gln ile
 SEQ ID NO: 281

SEQ ID NOS:280-281

FIG. 14B'

1/1 SEQ ID NO: 282   31/11

GTT TTC GAG TAC CGG CAG ATC CCA GGT TCA CCA GGT CTC ACC AGA TC
val phe glu tyr arg gln ile pro gly ser pro gly leu thr arg
 SEQ ID NO: 283

SEQ ID NOS:282-283

FIG. 14C'

ORF predicted based on the sequence   published by Cole et al.
(Nature 393:537-544) and containing seq14A'

```
1/1 ___ SEQ ID NO:  284                    31/11
TAG CGG TTC GGC GGG AAG CTA GCG GCG ACG  TTG TCG GTG GCC GGT GAT ATA TTG GGT CAG
AMB(arg phe gly gly lys leu ala ala thr  leu ser val ala gly asp ile leu gly gln
61/21 _____ SEQ ID NO:  285                91/31
ACG GGT ATG GCG GCG GCT GAG GTG ATC TGC  GAC ACG CCG CCG CGG TGC TCG AGC CAG GCT
thr gly met ala ala ala glu val ile cys  asp thr pro pro arg cys ser ser gln ala
121/41                                     151/51
TAC GAC CAG GGA ATT TCG AAA ATG TTA TTC  AGA ACA TCT TGT ATC TCT TCT CCG TGC CAC
tyr asp gln gly ile ser lys met leu phe  arg thr ser cys ile ser ser pro cys his
181/61                                     211/71
CCC CTA GGT GTA GTG TTT TCG AGT ACC GGC  AGA TCC CAG GTT CAC CAG GTC TCA CCA gat
pro leu gly val val phe ser ser thr gly  arg ser gln val his gln val ser pro asp
241/81                                     271/91
cca cgg ggc gcg atg aac ttc ccg gca tcg  gca tcg cca ggt cga cgg acg tgg tcg cgc
pro arg gly ala met asn phe pro ala ser  ala ser pro gly arg arg thr trp ser arg
301/101                                    331/111
tat gac ggg aat ctg gag cct tgt cgg gcc  gct caa cat atc gaa gat gca cta ctt gag
tyr asp gly asn leu glu pro cys arg ala  ala gln his ile glu asp ala leu leu glu
361/121                                    391/131
tcg ttg cca gat cct gtc aga ttc ccg att  tcc gca aag gag cgg tac gcc cat gac cgt
ser leu pro asp pro val arg phe pro ile  ser ala lys glu arg tyr ala his asp arg
421/141
gac cgt tta cac taa
asp arg leu his)OCH
```

SEQ ID NOS:284-285

FIG. 14D

Sequence Rv3054c predicted by Cole et al. (Nature 393:537-544)
which may be in phase with Seq14A'

1/1 _ **SEQ ID NO: 286**   31/11
gtg tca gat acc aag tcc gac atc aaa atc ttg gcc tta gtg gga agc ctg cgc gcg gcg
(val ser asp thr lys ser asp ile lys ile leu ala leu val gly ser leu arg ala ala
61/21 ~ **SEQ ID NO: 287**   91/31
tcg ttc aac cgc cag atc gcc gag ctg gct gcc aag gtc gct ccg gac ggc gtc acc gtc
ser phe asn arg gln ile ala glu leu ala ala lys val ala pro asp gly val thr val
121/41   151/51
acc atg ttc gag ggg ctg ggg gac ctg ccg ttc tac aac gaa gac atc gac aca gcg acg
thr met phe glu gly leu gly asp leu pro phe tyr asn glu asp ile asp thr ala thr
181/61   211/71
gag gtg ccg gcg ccg gtg agc gcg ttg cgg gag gcc gcg tct gac gcg cac gct gcc ttg
glu val pro ala pro val ser ala leu arg glu ala ala ser asp ala his ala ala leu
241/81   271/91
gtg gtc acg ccg gaa tac aac ggc agc att ccg gcc gtg atc aag aac gcg atc gac tgg
val val thr pro glu tyr asn gly ser ile pro ala val ile lys asn ala ile asp trp
301/101   331/111
ctg tcc agg cca ttc ggc gat ggc gcg ttg aag gac aag ccg ttg gcc gtg atc ggc ggc
leu ser arg pro phe gly asp gly ala leu lys asp lys pro leu ala val ile gly gly
361/121   391/131
tcc atg ggc cgc tac ggc ggg gta tgg gcg cac gac gag act cgc aag tcg ttc agc atc
ser met gly arg tyr gly gly val trp ala his asp glu thr arg lys ser phe ser ile
421/141   451/151
gct ggc acg cgg gtg gtc gat gcg atc aaa ctg tcg gtg ccg ttc caa act ctg ggc aag
ala gly thr arg val val asp ala ile lys leu ser val pro phe gln thr leu gly lys
481/161   511/171
tcg gtc gcg gac gac gcc ggg ctg gcg gcg aat gtg cgc gac gcc gtc ggc aac ttg gcc
ser val ala asp asp ala gly leu ala ala asn val arg asp ala val gly asn leu ala
541/181
gct gag gtc ggc tga
ala glu val gly)OPA

SEQ ID NOS:286-287

FIG. 14E

ORF predicted by Cole et al. (Nature 393:537-544) and containing Rv3054c

1/1 SEQ ID NO: 288   31/11
taa cgc gat cgg aat aaa tcg gac cat ggt ccg gtt ggc tcg tgc aag gac gtg gac caa
OCH(arg asp arg asn lys ser asp his gly pro val gly ser cys lys asp val asp gln
61/21 SEQ ID NO: 289   91/31
caa gcg gaa agg aac gta gca gtg tca gat acc aag tcc gac atc aaa atc ttg gcc tta
gln ala glu arg asn val ala val ser asp thr lys ser asp ile lys ile leu ala leu
121/41   151/51
gtg gga agc ctg cgc gcg gcg tcg ttc aac cgc cag atc gcc gag ctg gct gcc aag gtc
val gly ser leu arg ala ala ser phe asn arg gln ile ala glu leu ala ala lys val
181/61   211/71
gct ccg gac ggc gtc acc gtc acc atg ttc gag ggg ctg ggg gac ctg ccg ttc tac aac
ala pro asp gly val thr val thr met phe glu gly leu gly asp leu pro phe tyr asn
241/81   271/91
gaa gac atc gac aca gcg acg gag gtg ccg gcg ccg gtg agc gcg ttg cgg gag gcc gcg
glu asp ile asp thr ala thr glu val pro ala pro val ser ala leu arg glu ala ala
301/101   331/111
tct gac gcg cac gct gcc ttg gtg gtc acg ccg gaa tac aac ggc agc att ccg gcc gtg
ser asp ala his ala ala leu val val thr pro glu tyr asn gly ser ile pro ala val
361/121   391/131
atc aag aac gcg atc gac tgg ctg tcc agg cca ttc ggc gat ggc gcg ttg aag gac aag
ile lys asn ala ile asp trp leu ser arg pro phe gly asp gly ala leu lys asp lys
421/141   451/151
ccg ttg gcc gtg atc ggc ggc tcc atg ggc cgc tac ggc ggg gta tgg gcg cac gac gag
pro leu ala val ile gly gly ser met gly arg tyr gly gly val trp ala his asp glu
481/161   511/171
act cgc aag tcg ttc agc atc gct ggc acg cgg gtg gtc gat gcg atc aaa ctg tcg gtg
thr arg lys ser phe ser ile ala gly thr arg val val asp ala ile lys leu ser val
541/181   571/191
ccg ttc caa act ctg ggc aag tcg gtc gcg gac gac gcc ggg ctg gcg gcg aat gtg cgc
pro phe gln thr leu gly lys ser val ala asp asp ala gly leu ala ala asn val arg
601/201   631/211
gac gcc gtc ggc aac ttg gcc gct gag gtc ggc tga
asp ala val gly asn leu ala ala glu val gly)OPA

SEQ ID NOS:288-289

FIG. 14F

SEQ ID NO: 290: first frame nucleic acid
SEQ ID NO: 297: second frame nucleic acid
SEQ ID NO: 310: third frame nucleic acid
fragment based on the sequence published by Cole et al. (Nature 393:537-544)
and containing seq 14F' and seq 14P' SEQ ID NO: 291
1/1 SEQ ID NO: 290,297,310   31/11
taa cgc gat cgg aat aaa tcg gac cat ggt ccg gtt ggc tcg tgc aag gac gtg gac caa
OCH(arg asp arg asn lys ser asp his gly pro val gly ser cys lys asp val asp gln
(asn ala ile gly ile asn arg thr met val arg leu ala arg ala arg thr trp thr asn
(thr arg ser glu)OCH(ile gly pro trp ser gly trp leu val gln gly arg gly pro thr
61/21 SEQ ID NO: 298 SEQ ID NO: 311 91/31 SEQ ID NO: 312
caa gcg gaa agg aac gta gca gtg tca gat acc aag tcc gac atc aaa atc ttg gcc tta
gln ala glu arg asn val ala val ser asp thr lys ser asp ile lys ile leu ala leu
lys arg lys gly thr)AMB(gln cys gln ile pro ser pro thr ser lys ser trp pro)AMB
ser gly lys glu arg ser)ser val arg tyr gln val arg his gln asn leu gly leu ser
121/41   SEQ ID NO: 299   151/51
gtg gga agc ctg cgc gcg gcg tcg ttc aac cgc cag atc gcc gag ctg gct gcc aag gtc
val gly ser leu arg ala ala ser phe asn arg gln ile ala glu leu ala ala lys val
(trp glu ala cys ala arg arg arg ser thr ala arg ser pro ser trp leu pro arg ser
gly lys pro ala arg gly val val gln pro pro asp arg arg ala gly cys gln gly arg
181/61 SEQ ID NO: 300   211/71
gct ccg gac ggc gtc acc atg ttc gag ggg gac ctg ccg ttc tac aac
ala pro asp gly val thr val thr met phe glu gly leu gly asp leu pro phe tyr asn
leu arg thr ala ser pro ser pro cys ser arg gly trp gly thr cys arg ser thr thr
ser gly arg arg his arg his his val arg gly ala gly gly pro ala val leu gln arg
241/81   271/91
gaa gac atc gac aca gcg acg gag gtg ccg gcg ccg gtg agc gcg ttg cgg gag gcc gcg
glu asp ile asp thr ala thr glu val pro ala pro val ser ala leu arg glu ala ala
lys thr ser thr gln arg arg arg cys arg arg arg)OPA(ala arg cys gly arg pro arg
arg his arg his ser asp gly gly ala gly ala gly glu arg val ala gly gly arg val)
301/101   331/111   SEQ ID NO: 301
tct gac gcg cac gct gcc ttg gtg gtc acg ccg gaa tac aac ggc agc att ccg gcc gtg
ser asp ala his ala ala leu val val thr pro glu tyr asn gly ser ile pro ala val
leu thr arg thr leu pro trp trp ser arg arg asn thr thr ala ala phe arg pro OPA
OPA(arg ala arg cys leu gly gly his ala gly ile gln arg gln his ser gly arg asp
361/121 SEQ ID NO: 313   391/131
atc aag aac gcg atc gac tgg ctg tcc agg cca ttc ggc gat ggc gcg ttg aag gac aag
ile lys asn ala ile asp trp leu ser arg pro phe gly asp gly ala leu lys asp lys
(ser arg thr arg ser thr gly cys pro gly his ser ala met ala arg)OPA(arg thr ser
gln glu arg asp arg leu ala val gln ala ile arg arg trp arg val glu)gly gln ala
421/141 SEQ ID NO: 302   451/151   SEQ ID NO: 303
ccg ttg gcc gtg atc ggc ggc tcc atg ggc cgc tac ggc ggg gta tgg gcg cac gac gag
pro leu ala val ile gly gly ser met gly arg tyr gly gly val trp ala his asp glu
arg trp pro)OPA(ser ala ala pro trp ala ala thr ala gly tyr gly arg thr thr arg
val gly arg asp)arg arg leu his gly pro leu arg arg gly met gly ala arg arg asp
481/161   SEQ ID NO: 304   511/171
act cgc aag tcg ttc agc atc gct ggc acg cgg gtg gtc gat gcg atc aaa ctg tcg gtg
thr arg lys ser phe ser ile ala gly thr arg val val asp ala ile lys leu ser val
leu ala ser arg ser ala ser leu ala arg gly trp ser met arg ser asn cys arg cys
ser gln val val gln his arg trp his ala gly gly arg cys asp gln thr val gly ala
541/181   571/191
ccg ttc caa act ctg ggc aag tcg gtc gcg gac gac gcc ggg ctg gcg gcg aat gtg cgc
pro phe gln thr leu gly lys ser val ala asp asp ala gly leu ala ala asn val arg
arg ser lys leu trp ala ser arg ser arg thr thr pro gly trp arg arg met cys ala
val pro asn ser gly gln val gly arg gly arg arg arg ala gly gly glu cys ala arg
601/201   631/211   SEQ ID NO: 292
gac gcc gtc ggc aac ttg gcc gct gag gtc ggc tga tcc ctg ggc cga ggc ggg tca gcc
asp ala val gly asn leu ala ala glu val gly)OPA(ser leu gly arg gly gly ser ala
thr pro ser ala thr trp pro leu arg ser ala arg asp pro trp ala glu ala gly gln pro
arg arg arg gln leu gly arg)OPA(gly arg leu ile pro gly pro arg val ser gln)
661/221   SEQ ID NO: 314   691/231   SEQ ID NO: 293
aat agc ggc tcc atc ggc ttt gct ggt agc ggt tcg gcg gga agc tag cgg cga cgt tgt
asn ser gly ser ile gly phe ala gly ser gly ser ala gly ser)AMB(arg arg arg cys
ile ala ala pro ser ala leu leu val ala val arg arg glu ala ser gly asp val val
AMB(arg leu his arg leu cys trp)AMB(arg phe gly gly lys leu ala ala thr leu ser
SEQ ID NO: 315   SEQ ID NO: 316

SEQ ID NOS:290-316

FIG. 14G

721/241   751/251   **SEQ ID NO: 294**
cgg tgg ccg gtg ata tat tgg gtc aga cgg gta tgg cgg cgg ctg agg tga tct gcg aca
arg trp pro val ile tyr trp val arg arg val trp arg arg leu arg)OPA(ser ala thr
 gly gly arg)OPA(tyr ile gly ser asp gly tyr gly gly gly)OPA(gly asp leu arg his
   val ala gly/asp ile leu gly gln thr gly met ala ala ala/glu val ile cys asp thr
781/261   **SEQ ID NO: 305**   811/271   **SEQ ID NO: 306**
cgc cgc cgc ggt gct cga gcc agg ctt acg acc agg gaa ttt cga aaa tgt tat tca gaa
arg arg arg gly ala arg ala arg leu thr thr arg glu phe arg lys cys tyr ser glu
 ala ala ala val leu glu pro gly leu arg pro gly asn phe glu asn val ile gln asn
   pro pro arg cys ser ser gln ala tyr asp gln gly ile ser lys met leu phe arg thr
841/281   871/291   **SEQ ID NO: 295**
cat ctt gta tct ctt ctc cgt gcc acc ccc tag gtg tag tgt ttt cga gta ccg gca gat
his leu val ser leu leu arg ala thr pro)AMB val AMB(cys phe arg val pro ala asp
 ile leu tyr leu phe ser val pro pro pro arg cys ser val phe glu tyr arg gln ile
   ser cys ile ser ser pro cys his pro leu gly val val phe ser ser thr gly arg ser
901/301   931/311   **SEQ ID NO: 296**
ccc agg ttc acc agg tct cac cag atc cac ggg gcg cga tga act tcc cgg cat cgg cat
pro arg phe thr arg ser his gln ile his gly ala arg)OPA(thr ser arg his arg his
 pro gly ser pro gly leu thr arg ser thr gly arg asp glu leu pro gly ile gly ile
   gln val his gln val ser pro asp pro arg gly ala met asn phe pro ala ser ala ser
961/321   991/331
cgc cag gtc gac gga cgt ggt cgc gct atg acg gga atc tgg agc ctt gtc ggg ccg ctc
arg gln val asp gly arg gly arg ala met thr gly ile trp ser leu val gly pro leu
 ala arg ser thr asp val val ala leu)OPA(arg glu ser gly ala leu ser gly arg ser
   pro gly arg arg thr trp ser arg tyr/asp gly asn leu glu pro cys arg ala ala gln
1021/341   **SEQ ID NO: 307**   1051/351
aac ata tcg aag atg cac tac ttg agt cgt tgc cag atc ctg tca gat tcc cga ttt ccg
asn ile ser lys met his tyr leu ser arg cys gln ile leu ser asp ser arg phe pro
 thr tyr arg arg cys thr thr)OPA(val val ala arg ser cys gln ile pro asp phe arg
   his ile glu asp ala leu leu/glu ser leu pro asp pro val arg phe pro ile ser ala
1081/361   **SEQ ID NO: 308**   1111/371
caa agg agc ggt acg ccc atg acc gtg acc tac act aa)
gln arg ser gly thr pro met thr val thr val tyr thr)
 lys gly ala val arg pro)OPA pro OPA(pro phe thr leu
   lys glu arg tyr ala his asp arg asp)arg leu his)OCH
                SEQ ID NO: 309

                       SEQ ID NOS:290-316

FIG. 14G(continued)

1/1   **SEQ ID NO: 317**   31/11
CAA GCC CGG CCG CGA CTG TTT GCC GTT TTG GGG CTC CTA CCA GAA CAC CAC CTG GCG GCC
(gln ala arg pro arg leu phe ala val leu gly leu leu pro glu his his leu ala ala
61/21   **SEEQ ID NO: 318**   91/31
GCG CAC CAT GGT GTG CAC CAG TTG CGA TCG GTT CCT CCC GCG CGC GGG CGG CGA CGA CGT
ala his his gly val his gln leu arg ser val pro pro ala arg gly arg arg arg arg
121/41   151/51
CGA TGC CCG CGC CCC GGC GGC GCA GCT GCG TAG CTC GAC CCG GTC GAC GAC GAC GGG GTC
arg cys pro arg pro gly gly ala ala ala)AMB(leu asp pro val asp asp asp gly val
181/61   211/71   **SEQ ID NO: 319**
GGC GGA CCA GTC GGC GAT GTC GAG GCG ATG GCA ATA CAG CGC CTT GGT GCG CGG CCA CAC
gly gly pro val gly asp val glu ala met ala ile gln arg leu gly ala arg pro his
241/81   271/91
GTC TGA GGT GGC GAA GAC CAG TCC CGC GCC CAC CGG CAG CCG GAT CCG GAT ACG CGG TAC
val)OPA(gly gly glu asp gln ser arg ala his arg gln pro asp pro asp thr arg tyr)
            SEQ ID NO: 320

                       SEQ ID NOS:317-320

FIG. 15A

SEQ ID NO: 321   32/11
AAG CCC GGC CGC GAC TGT TTG CCG TTT TGG GGC TCC TAC CAG AAC ACC ACC TGG CGG CCG
lys pro gly arg asp cys leu pro phe trp gly ser tyr gln asn thr thr trp arg pro
62/21 ──── SEQ ID NO: 322   92/31
CGC ACC ATG GTG TGC ACC AGT TGC GAT CGG TTC CTC CCG CGC GCG GGC GGC GAC GAC GTC
arg thr met val cys thr ser cys asp arg phe leu pro arg ala gly gly asp asp val
122/41   152/51
GAT GCC CGC GCC CCG GCG GCG CAG CTG CGT AGC TCG ACC CGG TCG ACG ACG ACG GGG TCG
asp ala arg ala pro ala ala gln leu arg ser ser thr arg ser thr thr thr gly ser
182/61   212/71
GCG GAC CAG TCG GCG ATG TCG AGG CGA TGG CAA TAC AGC GCC TTG GTG CGC GGC CAC ACG
ala asp gln ser ala met ser arg arg trp gln tyr ser ala leu val arg gly his thr
242/81   272/91
TCT GAG GTG GCG AAG ACC AGT CCC GCG CCC ACC GGC AGC CGG ATC CGG ATA CGC GGT AC
·ser glu val ala lys thr ser pro ala pro thr gly ser arg ile arg ile arg gly)

SEQ ID NOS:321-322

FIG.15B

SEQ ID NO: 323   33/11
AGC CCG GCC GCG ACT GTT TGC CGT TTT GGG GCT CCT ACC AGA ACA CCA CCT GGC GGC CGC
ser pro ala ala thr val cys arg phe gly ala pro thr arg thr pro pro gly gly arg
63/21 ──── SEQ ID NO: 324   93/31
GCA CCA TGG TGT GCA CCA GTT GCG ATC GGT TCC TCC CGC GCG CGG GCG GCG ACG ACG TCG
ala pro trp cys ala pro val ala ile gly ser ser arg ala arg ala ala thr thr ser
123/41   153/51
ATG CCC GCG CCC CGG CGG CGC AGC TGC GTA GCT CGA CCC GGT CGA CGA CGA CGG GGT CGG
met pro ala pro arg arg arg ser cys val ala arg pro gly arg arg arg gly arg
183/61   213/71
CGG ACC AGT CGG CGA TGT CGA GGC GAT GGC AAT ACA GCG CCT TGG TGC GCG GCC ACA CGT
arg thr ser arg arg cys arg gly asp gly asn thr ala pro trp cys ala ala thr arg
243/81   273/91
CTG AGG TGG CGA AGA CCA GTC CCG CGC CCA CCG GCA GCC GGA TCC GGA TAC GCG GTA C
leu arg trp arg arg pro val pro arg pro pro ala ala gly ser gly tyr ala val)

SEQ ID NOS:323-324

FIG.15C

part of the nucleotide sequence of seq15A

```
1/1   SEQ ID NO: 325                        31/11
GGC GGC CGC GCG CCA TGG TGT GCA CCA GTT GCG ATC GGT TCT CCC GCG CGC GGG CGG CGA
gly gly arg ala pro trp cys ala pro val ala ile gly ser pro ala arg gly arg arg
61/21    SEQ ID NO: 326                     91/31
CGA CGT CGA TGG CCG CGC CCC GGC GGC TGC AGC TGC GTA GCT CGA CCC GGT CGA CGA CGA
arg arg arg trp pro arg pro gly gly cys ser cys val ala arg pro gly arg arg arg
121/41                                      151/51
CGG GGT CGG CGG GCC AGT CGG CGA TGT CGA GGC GAT GGC AAT ACA GCG CCT TGG TGC GCG
arg gly arg arg ala ser arg arg cys arg gly asp gly asn thr ala pro trp cys ala
181/61                                      211/71
GCC ACA CGT CTG AGG TGG CGA AGA CCA GTC CCG CGC CCA CCG GCA GCC GGA TC
ala thr arg leu arg trp arg arg pro val pro arg pro pro ala ala gly
```

SEQ ID NOS:325-326

FIG.15A'

```
1/1   SEQ ID NO: 327                        31/11
GCG GCC GCG CGC CAT GGT GTG CAC CAG TTG CGA TCG GTT CTC CCG CGC GCG GGC GGC GAC
(ala ala ala arg his gly val his gln leu arg ser val leu pro arg ala gly gly asp
61/21    SEQ ID NO: 328                     91/31
GAC GTC GAT GGC CGC GCC CCG GCG GCT GCA GCT GCG TAG CTC GAC CCG GTC GAC GAC GAC
asp val asp gly arg ala pro ala ala ala ala ala)AMB(leu asp pro val asp asp asp
121/41                                      151/51      SEQ ID NO: 329
GGG GTC GGC GGG CCA GTC GGC GAT GTC GAG GCG ATG GCA ATA CAG CGC CTT GGT GCG CGG
gly val gly gly pro val gly asp val glu ala met ala ile gln arg leu gly ala arg
181/61                                      211/71
CCA CAC GTC TGA GGT GGC GAA GAC CAG TCC CGC GCC CAC CGG CAG CCG GAT C
pro his val)OPA(gly gly glu asp gln ser arg ala his arg gln pro asp)
          SEQ ID NO: 330
```

SEQ ID NOS:327-330

FIG. 15B'

```
1/1   SEQ ID NO: 331                        31/11
TGG CGG CCG CGC GCC ATG GTG TGC ACC AGT TGC GAT CGG TTC TCC CGC GCG CGG GCG GCG
trp arg pro arg ala met val cys thr ser cys asp arg phe ser arg ala arg ala ala
61/21       SEQ ID NO: 332                  91/31
ACG ACG TCG ATG GCC GCG CCC CGG CGG CTG CAG CTG CGT AGC TCG ACC GGT CGA CGA CGA   (see note)
thr thr ser met ala ala pro arg arg leu gln leu arg ser ser thr arg ser thr thr
121/41                                      151/51
ACG GGG TCG GCG GGC CAG TCG GCG ATG TCG AGG CGA TGG CAA TAC AGC GCC TTG GTG CGC
thr gly ser ala gly gln ser ala met ser arg arg trp gln tyr ser ala leu val arg
181/61                                      211/71
GGC CAC ACG TCT GAG GTG GCG AAG ACC AGT CCC GCG CCC ACC GGC AGC CGG ATC
gly his thr ser glu val ala lys thr ser pro ala pro thr gly ser arg ile
```

SEQ ID NOS:331-332

FIG. 15C'

ORF containing Seq15A' according to Cole et al. (Nature 393:537-544)

```
1/1   ┌─ SEQ ID NO: 333                          31/11
taa ggt ccg cca acg ctt tac gct cga cgg ccg cca cga gtt ggc cgg cca ctt tca ggc
OCH(gly pro pro thr leu tyr ala arg arg pro pro arg val gly arg pro leu ser gly
61/21 ┕─ SEQ ID NO: 334                               91/31
cgt agt cgc cgc agg gca ggg ctt ccc gcg tcg tct tcg cgg gtt tgt cgg caa agg tgt
arg ser arg arg arg ala gly leu pro ala ser ser ser arg val cys arg gln arg cys
121/41                                          151/51
agg ggt agc gtt cgt ggg cgt cga cga cga tgt gca gct cgg gga tgc cgg cgg cgc ggg
arg gly ser val arg gly arg arg arg arg cys ala ala arg gly cys arg arg arg gly
181/61                                          211/71
cgg tgg ggg tgc gca cgc ccg gcc gcg act gtt tgc gcg ttt tgg ggc tct gcc aga aca
arg trp gly cys ala arg pro ala ala thr val cys ala phe trp gly ser ala arg thr
241/81                                          271/91
cca cct ggc ggc cgc gcg cca tgg tgt gca cca gtt gcg atc ggt tct ccc gcg cgc ggg
pro pro gly gly arg ala pro trp cys ala pro val ala ile gly ser pro ala arg gly
301/101                                         331/111
cgg cga cga cgt cga tgg ccg cgc ccc ggc ggc tgc agc tgc gta gct cga ccc ggt cga
arg arg arg arg arg trp pro arg pro gly gly cys ser cys val ala arg pro gly arg
361/121                                         391/131
cga cga cgg ggt cgg cgg gcc agt cgg cga tgt cga ggc gat ggc aat aca gcg cct tgg
arg arg arg gly arg arg ala ser arg arg cys arg gly asp gly asn thr ala pro trp
421/141                                         451/151
tgc gcg gcc aca cgt ctg agg tgg cga aga cca gtc ccg cgc cca ccg gca gcc gga tca
cys ala ala thr arg leu arg trp arg arg pro val pro arg pro pro ala ala gly ser
481/161                                         511/171
ggt agg gca ggc gcg agt ctt cag cgg ggt tgg cgg cga cga gca gct cca cag agt gtg
gly arg ala gly ala ser leu gln arg gly trp arg arg arg ala ala pro gln ser val
541/181                                         571/191
agg gta cgg gcg gcg tac ggc aac ggt gaa gca ggc act ccg acg aac cca tcg tca cgt
arg val arg ala ala tyr gly asn gly glu ala gly thr pro thr asn pro ser ser arg
601/201
cga agg ggc agg tga
arg arg gly arg) OPA
```

SEQ ID NOS:333-334

FIG. 15D

R:Rv2530c predicted according to Cole et al. (Nature 393:537-544) and which
may be in phase with SEQ15A
1/1   SEQ ID NO: 335   31/11
```
gtg acg gca ctg ctc gat gtc aat gtg ctg atc gcg ctg ggc tgg ccg aat cac gtt cac
(val thr ala leu leu asp val asn val leu ile ala leu gly trp pro asn his val his
```
61/21   SEQ ID NO: 336   91/31
```
cat gcg gcc gcg cag cga tgg ttc acg cag ttc tcc tcg aat ggg tgg gcc acc acg ccg
his ala ala ala gln arg trp phe thr gln phe ser ser asn gly trp ala thr thr pro
```
121/41   151/51
```
atc acc gag gca ggg tat gtc cga att tca agc aat cgc agt gtg atg cag gtg tcg acc
ile thr glu ala gly tyr val arg ile ser ser asn arg ser val met gln val ser thr
```
181/61   211/71
```
acg ccg gct atc gcg atc gct cag ttg gcg gcg atg act tct ctt gcc ggg cac acg ttt
thr pro ala ile ala ile ala gln leu ala ala met thr ser leu ala gly his thr phe
```
241/81   271/91
```
tgg cct gac gat gtg cca ctg atc gtt ggg agc gcc ggc gat cgc gat gcg gtg tcc aac
trp pro asp asp val pro leu ile val gly ser ala gly asp arg asp ala val ser asn
```
301/101   331/111
```
cac cgt cgg gtc acc gac tgc cat ctc atc gcc ttg gcc gcg cgc tac ggg ggc cgg ttg
his arg arg val thr asp cys his leu ile ala leu ala ala arg tyr gly gly arg leu
```
361/121   391/131
```
gtc aca ttc gat gcc gca ctg gcc gat tca gca tcc gca ggc ctc gtc gag gtg ttg tag
val thr phe asp ala ala leu ala asp ser ala ser ala gly leu val glu val leu)AMB
```

SEQ ID NOS:335-336

FIG. 15E

Seq15P: ORF according to Cole et al. (Nature 393:537-544) containing Rv2530c
1/1   SEQ ID NO: 337   31/11
```
tga tgt tcc gcc gga tgc gcc gac ggt gac ttc cga gga tgt cgt ccg cgc gct cga gga
OPA(cys ser ala gly cys ala asp gly asp phe arg gly cys arg pro arg ala arg gly
```
61/21   SEQ ID NO: 338   91/31
```
cga cgt gtg acg gca ctg ctc gat gtc aat gtg ctg atc gcg ctg ggc tgg ccg aat cac
arg arg val thr ala leu leu asp val asn val leu ile ala leu gly trp pro asn his
```
121/41   151/51
```
gtt cac cat gcg gcc gcg cag cga tgg ttc acg cag ttc tcc tcg aat ggg tgg gcc acc
val his his ala ala ala gln arg trp phe thr gln phe ser ser asn gly trp ala thr
```
181/61   211/71
```
acg ccg atc acc gag gca ggg tat gtc cga att tca agc aat cgc agt gtg atg cag gtg
thr pro ile thr glu ala gly tyr val arg ile ser ser asn arg ser val met gln val
```
241/81   271/91
```
tcg acc acg ccg gct atc gcg atc gct cag ttg gcg gcg atg act tct ctt gcc ggg cac
ser thr thr pro ala ile ala ile ala gln leu ala ala met thr ser leu ala gly his
```
301/101   331/111
```
acg ttt tgg cct gac gat gtg cca ctg atc gtt ggg agc gcc ggc gat cgc gat gcg gtg
thr phe trp pro asp asp val pro leu ile val gly ser ala gly asp arg asp ala val
```
361/121   391/131
```
tcc aac cac cgt cgg gtc acc gac tgc cat ctc atc gcc ttg gcc gcg cgc tac ggg ggc
ser asn his arg arg val thr asp cys his leu ile ala leu ala ala arg tyr gly gly
```
421/141   451/151
```
cgg ttg gtc aca ttc gat gcc gca ctg gcc gat tca gca tcc gca ggc ctc gtc gag gtg
arg leu val thr phe asp ala ala leu ala asp ser ala ser ala gly leu val glu val
```
481/161
```
ttg tag
leu)AMB
```

SEQ ID NOS:337-338

FIG. 15F

SEQ ID NO: 339: first frame nucleic acid
SEQ ID NO: 346: second frame nucleic acid
SEQ ID NO: 347: third frame nucleic acid
Fragment containing Seq15P' and Seq 15F'
1/1 _SEQ ID NO: 339,346,347   31/11
tga tgt tcc gcc gga tgc gcc gac ggt gac ttc cga gga tgt cgt ccg cgc gct cga gga
OPA(cys ser ala gly cys ala asp gly asp phe arg gly cys arg pro arg ala arg gly
 asp val pro pro asp ala pro thr val thr ser glu asp val val arg ala leu glu asp
   Met phe arg arg met arg arg arg)OPA(leu pro arg met ser ser ala arg ser arg thr
61/21 SEQ ID NO: 340 SEQ ID NO:348 91/31 SEQ ID NO: 349
cga cgt gtg acg gca ctg ctc gat gtc aat gtg ctg atc gcg ctg ggc tgg ccg aat cac
arg arg val thr ala leu leu asp val asn val leu ile ala leu gly trp pro asn his
 asp val OPA arg his cys ser met ser met cys OPA ser arg trp ala gly arg ile thr
   thr cys asp gly thr ala arg cys gln cys ala asp arg ala gly leu ala glu ser arg
121/41   151/51
gtt cac cat gcg gcc gcg cag cga tgg ttc acg cag ttc tcc tcg aat ggg tgg gcc acc
val his his ala ala ala gln arg trp phe thr gln phe ser ser asn gly trp ala thr
 phe thr met arg pro arg ser asp gly ser arg ser ser pro arg met gly gly pro pro
   ser pro cys gly arg ala ala met val his ala val leu leu glu trp val gly his his
181/61   211/71
acg ccg atc acc gag gca ggg tat gtc cga att tca agc aat cgc agt gtg atg cag gtg
thr pro ile thr glu ala gly tyr val arg ile ser ser asn arg ser val met gln val
 arg arg ser pro arg gln gly met ser glu phe gln ala ile ala val OPA cys arg cys
   ala asp his arg gly arg val cys pro asn phe lys gln ser gln cys asp ala gly val
241/81   271/91
tcg acc acg ccg gct atc gcg atc gct cag ttg gcg gcg atg act tct ctt gcc ggg cac
ser thr thr pro ala ile ala ile ala gln leu ala ala met thr ser leu ala gly his
 arg pro arg arg leu ser arg ser leu ser trp arg arg OPA leu leu leu pro gly thr
   asp his ala gly tyr arg asp arg ser val gly gly asp asp phe ser cys arg ala his
301/101   331/111
acg ttt tgg cct gac gat gtg cca ctg atc gtt ggg agc gcc ggc gat cgc gat gcg gtg
thr phe trp pro asp asp val pro leu ile val gly ser ala gly asp arg asp ala val
 arg phe gly leu thr met cys his OPA ser leu gly ala pro ala ile ala met arg cys
   val leu ala)OPA(arg cys ala thr asp arg trp glu arg arg arg ser arg cys gly val
361/121   SEQ ID NO: 350   391/131
tcc aac cac cgt cgg gtc acc gac tgc cat ctc atc gcc ttg gcc gcg cgc tac ggg ggc
ser asn his arg arg val thr asp cys his leu ile ala leu ala ala arg tyr gly gly
 pro thr thr val gly ser pro thr ala ile ser ser pro trp pro arg ala thr gly ala
   gln pro pro ser gly his arg leu pro ser his arg leu gly arg ala leu arg gly pro
421/141   451/151
cgg ttg gtc aca ttc gat gcc gca ctg gcc gat tca gca tcc gca ggc ctc gtc gag gtg
arg leu val thr phe asp ala ala leu ala asp ser ala ser ala gly leu val glu val
 gly trp ser his ser met pro his trp pro ile gln his pro gln ala ser ser arg cys
   val gly his ile arg cys arg thr gly arg phe ser ile arg arg pro arg arg gly val
481/161   SEQ ID NO: 341   511/171
ttg tag tca ccg/ggg atg ggc ggc tcg cca ggc ctg cag gat ctg cgg gcg cag gcg ccc
leu)AMB(ser pro gly met gly gly ser pro gly leu gln asp leu arg ala gln ala pro
 cys ser his arg gly trp ala ala arg gln ala cys arg ile cys gly arg arg arg pro
   val val thr gly asp gly arg leu ala arg pro ala gly ser ala gly ala gly ala pro
541/181   571/191
ccg gtc gga cac cgg cag gcc gac gct ttt ggc cca cgc gcg cag ctc ggc gct gct ggg
pro val gly his arg gln ala asp ala phe gly pro arg ala gln leu gly ala ala gly
 arg ser asp thr gly arg pro thr leu leu ala his ala arg ser ser ala leu leu gly
   gly arg thr pro ala gly arg arg phe trp pro thr arg ala ala arg arg cys trp ala
601/201   631/211
ctc ggg ctc ggc ggc agc cgg ctc gaa aac cgt ggt ggc gtc ggc atc gtc gac gaa cca
leu gly leu gly gly ser arg leu glu asn arg gly gly val gly ile val asp glu pro
 ser gly ser ala ala ala gly ser lys thr val val ala ser ala ser ser thr asn gln
   arg ala arg arg gln pro ala arg lys pro trp trp arg arg his arg arg arg thr arg)

SEQ ID NOS: 339-352

FIG. 15G

661/221   **SEQ ID NO: 342**   691/231
ggt gag ggc ggc ggc tag}ata gcg gta ggt gta ttc ctg ggc gag ctt gcg ggt ttg gca
gly glu gly gly gly)AMB(ile ala val gly val phe leu gly glu leu ala gly leu ala
 val arg ala ala ala arg AMB arg AMB val tyr ser trp ala ser leu arg val trp gln
  OPA(gly arg arg leu asp ser gly arg cys ile pro gly arg ala cys gly phe gly arg
721/241⎯SEQ ID NO: 351   751/251
gaa cac gat cgg cac gtt ggg aaa gcc gat ctg caa ttc ggc cag ccc atc ggc gat cgc
glu his asp arg his val gly lys ala asp leu gln phe gly gln pro ile gly asp arg
 asn thr ile gly thr leu gly lys pro ile cys asn ser ala ser pro ser ala ile ala
  thr arg ser ala arg trp glu ser arg ser ala ile arg pro ala his arg arg ser pro
781/261   811/271
cgt cgg gcg ggc gaa gga gtg cgc gaa gat ctc cga gta gcg gtc ctc gac cac cac ggc
arg arg ala gly glu gly val arg glu asp leu arg val ala val leu asp his his gly
 val gly arg ala lys glu cys ala lys ile ser glu AMB arg ser ser thr thr thr ala
  ser gly gly arg arg ser ala arg arg ser pro ser ser gly pro arg pro pro arg arg
841/281   871/291
ggc ccg tgg cag cgc ggc cag ttc ggt cag ttg gta ttt cag gtt gcc gtt cag cac gcc
gly pro trp gln arg gly gln phe gly gln leu val phe gln val ala val gln his ala
 ala arg gly ser ala ala ser ser val ser trp tyr phe arg leu pro phe ser thr pro
  pro val ala ala arg pro val arg ser val gly ile ser gly cys arg ser ala arg gln
901/301   931/311
aga agt aag gtc cgc caa cgc ttt acg ctc gac ggc cgc cac gag ttg gcc ggc cac ttt
arg ser lys val arg gln arg phe thr leu asp gly arg his glu leu ala gly his phe
 glu val arg ser ala asn ala leu arg ser thr ala ala thr ser trp pro ala thr phe
  lys)OCH(gly pro pro thr leu tyr ala arg arg pro pro arg val gly arg pro leu ser
961/321   ⟍SEQ ID NO. 352   991/331
cag gcc gta gtc gcc gca ggg cag ggc ttc ccg cgt cgt ctt cgc ggg ttt gtc ggc aaa
gln ala val val ala ala gly gln gly phe pro arg arg leu arg gly phe val gly lys
 arg pro AMB ser pro gln gly arg ala ser arg val val phe ala gly leu ser ala lys
  gly arg ser arg arg arg ala gly leu pro ala ser ser ser arg val cys arg gln arg
1021/341   1051/351
ggt gta ggg gta gcg ttc gtg ggc gtc gac gac gat gtg cag ctc ggg gat gcc ggc ggc .
gly val gly val ala phe val gly val asp asp asp val gln leu gly asp ala gly gly
 val AMB gly AMB arg ser trp ala ser thr thr met cys ser ser gly met pro ala ala
  cys arg gly ser val arg gly arg arg arg arg cys ala ala arg gly cys arg arg arg
1081/361   1111/371
gcg ggc ggt ggg ggt gcg cac gcc cgg ccg cga ctg ttt gcg cgt ttt ggg gct ctg cca
ala gly gly gly gly ala his ala arg pro arg leu phe ala arg phe gly ala leu pro
 arg ala val gly val arg thr pro gly arg asp cys leu arg val leu gly leu cys gln
  gly arg trp gly cys ala arg pro ala ala thr val cys ala phe trp gly ser ala arg
1141/381   1171/391
gaa cac cac ctg gcg gcc gcg cgc cat ggt gtg cac cag ttg cga tcg gtt ctc ccg cgc
glu his his leu ala ala ala arg his gly val his gln leu arg ser val leu pro arg
 asn thr thr trp arg pro arg ala met val cys thr ser cys asp arg phe ser arg ala
  thr pro pro gly gly arg ala pro trp cys ala pro val ala ile gly ser pro ala arg
1201/401   1231/411   **SEQ ID NO: 343**⟍
gcg ggc ggc gac gac gtc gat ggc cgc gcc ccg gcg gct gca gct gcg tag⟍ctc gac ccg
ala gly gly asp asp val asp gly arg ala pro ala ala ala ala ala)AMB(leu asp pro
 arg ala ala thr thr ser met ala ala pro arg arg leu gln leu arg ser ser thr arg
  gly arg arg arg arg arg trp pro arg pro gly cys ser cys val ala arg pro gly
1261/421   1291/431
gtc gac gac gac ggg gtc ggc ggg cca gtc ggc gat gtc gag gcg atg gca ata cag cgc
val asp asp asp gly val gly gly pro val gly asp val glu ala met ala ile gln arg
 ser thr thr thr gly ser ala gly gln ser ala met ser arg arg trp gln tyr ser ala
  arg arg arg arg gly arg arg ala ser arg arg cys arg gly asp gly asn thr ala pro

SEQ ID NOS:339-352

FIG. 15G (continued 1)

```
1321/441        SEQ ID NO: 344          1351/451
ctt ggt gcg cgg cca cac gtc tga ggt ggc gaa gac cag tcc cgc gcc cac cgg cag ccg
leu gly ala arg pro his val)OPA(gly gly glu asp gln ser arg ala his arg gln pro
 leu val arg gly his thr ser glu val ala lys thr ser pro ala pro thr gly ser arg
  trp cys ala ala thr arg leu arg trp arg arg pro val pro arg pro pro ala ala gly
1381/461                                1411/471
gat cag gta ggg cag gcg cga gtc ttc agc ggg gtt ggc ggc gac gag cag ctc cac aga
asp gln val gly gln ala arg val phe ser gly val gly gly asp glu gln leu his arg
 ile arg AMB gly arg arg glu ser ser ala gly leu ala ala thr ser ser ser thr glu
  ser gly arg ala gly ala ser leu gln arg gly trp arg arg arg ala ala pro gln ser
1441/481     SEQ ID NO: 345            1471/491
gtg tga ggg tac ggg cgg cgt acg gca acg gtg aag cag gca ctc cga cga acc cat cgt
val)OPA(gly tyr gly arg arg thr ala thr val lys gln ala leu arg arg thr his arg
 cys glu gly thr gly gly val arg gln arg OPA ser arg his ser asp glu pro ile val
  val arg val arg ala ala tyr gly asn gly glu ala gly thr pro thr asn pro ser ser
1501/501
cac gtc gaa ggg gca ggt ga)
his val glu gly ala gly)
 thr ser lys gly gln val
  arg arg arg gly arg)OPA
```

FIG. 15G (continued (2)

```
   SEQ ID NO: 353                       31/11
TGC GCA TGC CGA CCA GTG TGG TTG GCC GGA GTT CGT TTG TTC GCG ATT GCC TCA ACG ATT
(cys ala cys arg pro val trp leu ala gly val arg leu phe ala ile ala ser thr ile
61/21      SEQ ID NO: 354              91/31
CGA TAT AAC CAC TCT AGT CAC ATC AAC CAC ACT CGT ACC ATC GAG CGT GTG GGT TCA TGC
arg tyr asn his ser ser his ile asn his thr arg thr ile glu arg val gly ser cys
121/41                                  151/51
CAT GCA TTC GCG ACC GCG GGA GCC GGC GAA CCC GGC GCC ACA CAT AAT CCA GAT TGA GGA
his ala phe ala thr ala gly ala gly glu pro gly ala thr his asn pro asp)OPA(gly
181/61                                  211/71           SEQ ID NO: 355
GAC TTC CGT GCC GAA CCG ACG CCG ACG CAA GCT TTC GAC AGC CAT GAG CGC GGT CGC CGC
asp phe arg ala glu pro thr pro thr gln ala phe asp ser his glu arg gly arg arg
241/81                                  271/91
CCT GGC AGT TGC AAG TCC TTG TGC ATA TTT TCT TGT CTA CGA ATC AAC CGA AAC GAC CGA
pro gly ser cys lys ser leu cys ile phe ser cys leu arg ile asn arg asn asp arg
301/101                                 331/111
GCG GCC CGA GCA CCA TGA ATT CAA GCA GGC GGC GGT GTT GAC CGA CCT GCC CGG CGA GCT
ala ala arg ala pro)OPA(ile gln ala gly gly gly val asp arg pro ala arg arg ala
361/12   SEQ ID NO: 356                391/131
GAT GTC CGC GCT ATC GCA GGG GTT GTC CCA GTT CGG GAT C
asp val arg ala ile ala gly val val pro val arg asp)
```

SEQ ID NOS:353-356

FIG. 16A

SEQ ID NO: 357   32/11
GCG CAT GCC GAC CAG TGT GGT TGG CCG GAG TTC GTT TGT TCG CGA TTG CCT CAA CGA TTC
(ala his ala asp gln cys gly trp pro glu phe val cys ser arg leu pro gln arg phe
62/21   SEQ ID NO: 358   92/31
GAT ATA ACC ACT CTA GTC ACA TCA ACC ACA CTC GTA CCA TCG AGC GTG TGG GTT CAT GCC
asp ile thr thr leu val thr ser thr thr leu val pro ser ser val trp val his ala
122/41   152/51
ATG CAT TCG CGA CCG CGG GAG CCG GCG AAC CCG GCG CCA CAC ATA ATC CAG ATT GAG GAG
met his ser arg pro arg glu pro ala asn pro ala pro his ile ile gln ile glu glu
182/61   212/71
ACT TCC GTG CCG AAC CGA CGC CGA CGC AAG CTT TCG ACA GCC ATG AGC GCG GTC GCC GCC
thr ser val pro asn arg arg arg arg lys leu ser thr ala met ser ala val ala ala
242/81   272/91
CTG GCA GTT GGA AGT CCT TGT GCA TAT TTT CTT GTC TAC GAA TCA ACC GAA ACG ACC GAG
leu ala val ala ser pro cys ala tyr phe leu val tyr glu ser thr glu thr thr glu
302/101   332/111
CGG CCC GAG CAC CAT GAA TTC AAG CAG GCG GCG GTG TTG ACC GAC CTG CCC GGC GAG CTG
arg pro glu his his glu phe lys gln ala ala val leu thr asp leu pro gly glu leu
362/121   392/131
ATG TCC GCG CTA TCG CAG GGG TTG TCC CAG TTC GGG ATC
met ser ala leu ser gln gly leu ser gln phe gly ile)

SEQ ID NOS:357-358

FIG. 16B

SEQ ID NO: 359   33/11
CGC ATG CCG ACC AGT GTG GTT GGC CGG AGT TCG TTT GTT CGC GAT TGC CTC AAC GAT TCG
(arg met pro thr ser val val gly arg ser ser phe val arg asp cys leu asn asp ser
63/21   SEQ ID NO: 360   93/31
ATA TAA CCA CTC TAG TCA CAT CAA CCA CAC TCG TAC CAT CGA GCG TGT GGG TTC ATG CCA
ile)OCH pro leu AMB(ser his gln pro his ser tyr his arg ala cys gly phe met pro
123/41   SEQ ID NO: 926   153/51
TGC ATT CGC GAC CGC GGG AGC CGG CGA ACC CGG CGC CAC ACA TAA TCC AGA TTG AGG AGA
cys ile arg asp arg gly ser arg arg thr arg arg his thr)OCH(ser arg leu arg arg
183/61   213/71   SEQ ID NO: 927
CTT CCG TGC CGA ACC GAC GCC GAC GCA AGC TTT CGA CAG CCA TGA GCG CGG TCG CCG CCC
leu pro cys arg thr asp ala asp ala ser phe arg gln pro)OPA(ala arg ser pro pro
243/81   273/91   SEQ ID NO: 928
TGG CAG TTG CAA GTC CTT GTG CAT ATT TTC TTG TCT ACG AAT CAA CCG AAA CGA CCG AGC
trp gln leu gln val leu val his ile phe leu ser thr asn gln pro lys arg pro ser
303/101   333/111
GGC CCG AGC ACC ATG AAT TCA AGC AGG CGG CGG TGT TGA CCG ACC TGC CCG GCG AGC TGA
gly pro ser thr met asn ser ser arg arg arg cys)OPA(pro thr cys pro ala ser)OPA
363/121   393/131   SEQ ID NO: 929
TGT CCG CGC TAT CGC AGG GGT TGT CCC AGT TCG GGA TC
(cys pro arg tyr arg arg gly cys pro ser ser gly)
      SEQ ID NO: 930

SEQ ID NOS:359-360,926-930

FIG. 16C

~SEQ ID NO: 361   31/11
GCG GGC CAC CGA TCA GTC GAT CGG GTG GTT TCC GCT CCA TCA GCC CGG AAT TGA GGT GCC
(ala gly his arg ser val asp arg val val ser ala pro ser ala arg asn)OPA(gly ala
61/21   ~SEQ ID NO: 362   91/31   SEQ ID NO: 363
GCA GTG ACG ACA CCA GCG CAG GAC GCG CCG TTG GTG TTT CCC TCT GTT GCT TTC CCG TCC
ala val thr thr pro ala gln asp ala pro leu val phe pro ser val ala phe pro ser
121/41   151/51
GGC TCG CCT TTT TTT CAT CAA CGT TGG ACT GCC GCA GTG GCG ATG TTG GTC GCC GGC GTG
gly ser pro phe phe his gln arg trp thr ala ala val ala met leu val ala gly val
181/61   211/71
TTC GGT CAC CTG ACG GTC GGG ATG TTC CTT GGG TCT CGG GTT GCT GCT GGG TTT GCT CAA
phe gly his leu thr val gly met phe leu gly ser arg val ala ala gly phe ala gln
241/81   271/91
TGC CCT GCT GGT GCG GCG TTC GGC CGA GTC GAT CAC CGC CAA AGA GCA CCC GTT AAA ACG
cys pro ala gly ala ala phe gly arg val asp his arg gln arg ala pro val lys thr
301/101   331/111
GTC GAT GGC CCT CAA CTC GGC ATC GCG ACT GGC GAT TAT CAC CAT GCC TCG GGC TGA TC
val asp gly pro gln leu gly ile ala thr gly asp tyr his his ala ser gly)OPA

SEQ ID NOS:361-363

FIG. 17A

~SEQ ID NO: 364   32/11
CGG GCC ACC GAT CAG TCG ATC GGG TGG TTT CCG CTC CAT CAG CCC GGA ATT GAG GTG CCG
(arg ala thr asp gln ser ile gly trp phe pro leu his gln pro gly ile glu val pro
62/21   ~SEQ ID NO: 365   92/31
CAG TGA CGA CAC CAG CGC AGG ACG CGC CGT TGG TGT TTC CCT CTG TTG CTT TCC CGT CCG
gln)OPA(arg his gln arg arg thr arg arg trp cys phe pro leu leu leu ser arg pro
122/41   ~SEQ ID NO: 366   152/51
GCT CGC CTT TTT TTC ATC AAC GTT GGA CTG CCG CAG TGG CGA TGT TGG TCG CCG GCG TGT
ala arg leu phe phe ile asn val gly leu pro gln trp arg cys trp ser pro ala cys
182/61   212/71
TCG GTC ACC TGA CGG TCG GGA TGT TCC TTG GGT CTC GGG TTG CTG CTG GGT TTG CTC AAT
ser val thr)OPA(arg ser gly cys ser leu gly leu gly leu leu leu gly leu leu asn
242/81   ~SEQ ID NO: 367   272/91
GCC CTG CTG GTG CGG CGT TCG GCC GAG TCG ATC ACC GCC AAA GAG CAC CCG TTA AAA CGG
ala leu leu val arg arg ser ala glu ser ile thr ala lys glu his pro leu lys arg
302/101   332/111
TCG ATG GCC CTC AAC TCG GCA TCG CGA CTG GCG ATT ATC ACC ATG CCT CGG GCT GAT C
ser met ala leu asn ser ala ser arg leu ala ile ile thr met pro arg ala asp)

SEQ ID NOS:364-367

FIG. 17B

SEQ ID NO: 368   33/11
GGG CCA CCG ATC AGT CGA TCG GGT GGT TTC CGC TCC ATC AGC CCG GAA TTG AGG TGC CGC
(gly pro pro ile ser arg ser gly gly phe arg ser ile ser pro glu leu arg cys arg
63/21   SEQ ID NO: 369   93/31
AGT GAC GAC ACC AGC GCA GGA CGC GCC GTT GGT GTT TCC CTC TGT TGC TTT CCC GTC CGG
ser asp asp thr ser ala gly arg ala val gly val ser leu cys cys phe pro val arg
123/41   153/51
CTC GCC TTT TTT TCA TCA ACG TTG GAC TGC CGC AGT GGC GAT GTT GGT CGC CGG CGT GTT
leu ala phe phe ser ser thr leu asp cys arg ser gly asp val gly arg arg arg val
183/61   213/71
CGG TCA CCT GAC GGT CGG GAT GTT CCT TGG GTC TCG GGT TGC TGC TGG GTT TGC TCA ATG
arg ser pro asp gly arg asp val pro trp val ser gly cys cys trp val cys ser met
243/81   273/91
CCC TGC TGG TGC GGC GTT CGG CCG AGT CGA TCA CCG CCA AAG AGC ACC CGT TAA AAC GGT
pro cys trp cys gly val arg pro ser arg ser pro pro lys ser thr arg)OCH(asn gly
303/101   333/111   SEQ ID NO: 370
CGA TGG CCC TCA ACT CGG CAT CGC GAC TGG CGA TTA TCA CCA TGC CTC GGG CTG ATC
arg trp pro ser thr arg his arg asp trp arg leu ser pro cys leu gly leu ile)

SEQ ID NOS:368-370

FIG. 17C

part of the nucleotide sequence of seq17A

1/1   SEQ ID NO: 371   31/11
(ggc tag aac ccc gaa gga gac ctc gcg ggt tgc cgg ccc ccg gcc cat cgg atg cgt atc
gly AMB(asn pro glu gly asp leu ala gly cys arg pro pro ala his arg met arg ile
61/21   SEQ ID NO: 372   91/31
cgg tcg cgc cga ttc acg acc gac ata ggg agc tac ccc ttg ggt gat tcc ggt gcg acg
arg ser arg arg phe thr thr asp ile gly ser tyr pro leu gly asp ser gly ala thr
121/41   151/51
act gcg ata cgc tcg gcg ggc cac cga tca gtc gat cgg gtg gtt tcc gct cca tca gcc
thr ala ile arg ser ala gly his arg ser val asp arg val val ser ala pro ser ala
181/61   211/71
cgg aat tga ggt gcc gca gtg acg aca cca gcg cag gac gcg ccg ttg gtg ttt ccc tct
arg asn)OPA(gly ala ala val thr thr pro ala gln asp ala pro leu val phe pro ser
241/81   SEQ ID NO: 373   271/91
gtt gct ttc cgt ccg gtt cgc ctt ttt ttc atc aac gtt gga ctg gcc gca gtg gcg atg
val ala phe arg pro val arg leu phe phe ile asn val gly leu ala ala val ala met
301/101   331/111
ttg gtc gcc ggc gtg ttc ggt cac ctg acg gtc ggg atg ttc ttg ggt ctc ggg ttg ctg
leu val ala gly val phe gly his leu thr val gly met phe leu gly leu gly leu leu
361/121   391/131
ctg ggt ttg ctc aat gcc ctg ctg gtg cgg cgt tcg gcc gag tcg atc acc gcc aaa gag
leu gly leu leu asn ala leu leu val arg arg ser ala glu ser ile thr ala lys glu
421/141   451/151
cac ccg tta aaa cgg tcg atg gcc ctc aac tcg gca tcg cga ctg gcg att atc acc atc
his pro leu lys arg ser met ala leu asn ser ala ser arg leu ala ile ile thr ile
481/161
ctc ggg ctg atc
leu gly leu ile)

SEQ ID NOS:371-373

FIG. 17A'

```
1/1  SEQ ID NO: 374                      31/11
gct aga acc ccg aag gag acc tcg cgg gtt gcc ggc ccc cgg ccc atc gga tgc gta tcc
(ala arg thr pro lys glu thr ser arg val ala gly pro arg pro ile gly cys val ser
61/21  SEQ ID NO: 375                    91/31
ggt cgc gcc gat tca cga ccg aca tag gga gct acc cct tgg gtg att ccg gtg cga cga
gly arg ala asp ser arg pro thr)AMB(gly ala thr pro trp val ile pro val arg arg
121/41              SEQ ID NO: 376       151/51
ctg cga tac gct cgg cgg gcc acc gat cag tcg atc ggg tgg ttt ccg ctc cat cag ccc
leu arg tyr ala arg arg ala thr asp gln ser ile gly trp phe pro leu his gln pro
181/61                                   211/71
gga att gag gtg ccg cag tga cga cac cag cgc agg acg cgc cgt tgg tgt ttc cct ctg
gly ile glu val pro gln)OPA(arg his gln arg arg thr arg arg trp cys phe pro leu
241/81      SEQ ID NO: 377               271/91
ttg ctt tcc gtc cgg ttc gcc ttt ttt tca tca acg ttg gac tgg ccg cag tgg cga tgt
leu leu ser val arg phe ala phe phe ser ser thr leu asp trp pro gln trp arg cys
301/101                                  331/111
tgg tcg ccg gcg tgt tcg gtc acc tga cgg tcg gga tgt tct tgg gtc tcg ggt tgc tgc
trp ser pro ala cys ser val thr)OPA(arg ser gly cys ser trp val ser gly cys cys
361/121             SEQ ID NO: 378       391/131
tgg gtt tgc tca atg ccc tgc tgg tgc ggc gtt cgg ccg agt cga tca ccg cca aag agc
trp val cys ser met pro cys trp cys gly val arg pro ser arg ser pro pro lys ser
421/141                                  451/151
acc cgt taa aac ggt cga tgg ccc tca act cgg cat cgc gac tgg cga tta tca cca tcc
thr arg)OCH(asn gly arg trp pro ser thr arg his arg asp trp arg leu ser pro ser
481/161        SEQ ID NO: 379
tcg ggc tga tc
ser gly)OPA
```

SEQ ID NOS:374-379

FIG. 17B'

```
1/1   SEQ ID NO: 380                     31/11
cta gaa ccc cga agg aga cct cgc ggg ttg ccg gcc ccc ggc cca tcg gat gcg tat ccg
(leu glu pro arg arg arg pro arg gly leu pro ala pro gly pro ser asp ala tyr pro
61/21  SEQ ID NO: 381                    91/31
gtc gcg ccg att cac gac cga cat agg gag cta ccc ctt ggg tga ttc cgg tgc gac gac
val ala pro ile his asp arg his arg glu leu pro leu gly)OPA(phe arg cys asp asp
121/41                                   151/51            SEQ ID NO: 382
tgc gat acg ctc ggc ggg cca ccg atc agt cga tcg gt ggt ttc cgc tcc atc agc ccg
cys asp thr leu gly gly pro pro ile ser arg ser gly gly phe arg ser ile ser pro
181/61                                   211/71
gaa ttg agg tgc cgc agt gac gac acc agc gca gga cgc gcc gtt ggt gtt tcc ctc tgt
glu leu arg cys arg ser asp asp thr ser ala gly arg ala val gly val ser leu cys
241/81                                   271/91
tgc ttt ccg tcc ggt tcg cct ttt ttt cat caa cgt tgg act ggc cgc agt ggc gat gtt
cys phe pro ser gly ser pro phe phe his gln arg trp thr gly arg ser gly asp val
301/101                                  331/111
ggt cgc cgg cgt gtt cgg tca cct gac ggt cgg gat gtt ctt ggg tct cgg gtt gct gct
gly arg arg arg val arg ser pro asp gly arg asp val leu gly ser arg val ala ala
361/121                                  391/131
ggg ttt gct caa tgc cct gct ggt gcg gcg ttc ggc cga gtc gat cac cgc caa aga gca
gly phe ala gln cys pro ala gly ala ala phe gly arg val asp his arg gln arg ala
421/141                                  451/151
ccc gtt aaa acg gtc gat ggc cct caa ctc ggc atc gcg act ggc gat tat cac cat cct
pro val lys thr val asp gly pro gln leu gly ile ala thr gly asp tyr his his pro
481/161
cgg gct gat c
arg ala asp)
```

SEQ ID NOS:380-382

FIG. 17C'

sequence Rv1303 predicted by Cole et al. (Nature 393:537-544) and partially
containing Seq17A'

1/1 —SEQ ID NO: 383   31/11
atg acg aca cca gcg cag gac gcg ccg ttg gtg ttt ccc tct gtt gct ttc cgt ccg gtt
(met thr thr pro ala gln asp ala pro leu val phe pro ser val ala phe arg pro val
61/21 —SEQ ID NO: 384   91/31
cgc ctt ttt ttc atc aac gtt gga ctg gcc gca gtg gcg atg ttg gtc gcc ggc gtg ttc
arg leu phe phe ile asn val gly leu ala ala val ala met leu val ala gly val phe
121/41   151/51
ggt cac ctg acg gtc ggg atg ttc ttg ggt ctc ggg ttg ctg ctg ggt ttg ctc aat gcc
gly his leu thr val gly met phe leu gly leu gly leu leu leu gly leu leu asn ala
181/61   211/71
ctg ctg gtg cgg cgt tcg gcc gag tcg atc acc gcc aaa gag cac ccg tta aaa cgg tcg
leu leu val arg arg ser ala glu ser ile thr ala lys glu his pro leu lys arg ser
241/81   271/91
atg gcc ctc aac tcg gca tcg cga ctg gcg att atc acc atc ctc ggg ctg atc atc gcc
met ala leu asn ser ala ser arg leu ala ile ile thr ile leu gly leu ile ile ala
301/101   331/111
tac att ttc cgg ccc gct gga ttg ggc gtc gtg ttc ggg ctg gca ttc ttc cag gtg ctg
tyr ile phe arg pro ala gly leu gly val val phe gly leu ala phe phe gln val leu
361/121   391/131
ctg gtg gca acg acg gcc ctg ccg gtc ctg aag aag ctg cgc act gcg acc gag gaa ccg
leu val ala thr thr ala leu pro val leu lys lys leu arg thr ala thr glu glu pro
421/141   451/151
gtc gca act tat tct tcc aat ggc cag acc ggg gga tcg gaa gga agg agc gcc agc gat
val ala thr tyr ser ser asn gly gln thr gly gly ser glu gly arg ser ala ser asp
481/161
gac tga
asp)OPA

SEQ ID NOS:383-384

FIG. 17D

Orf according to Cole et al. (Nature 393:537-544) and containing Rv1303

1/1 —SEQ ID NO: 385   31/11
tga ggt gcc gac gtg acg aca cca gcg cag gac gcg ccg ttg gtg ttt ccc tct gtt gct
OPA(gly ala ala val thr thr pro ala gln asp ala pro leu val phe pro ser val ala
61/21 —SEQ ID NO: 386   91/31
ttc cgt ccg gtt cgc ctt ttt ttc atc aac gtt gga ctg gcc gca gtg gcg atg ttg gtc
phe arg pro val arg leu phe phe ile asn val gly leu ala ala val ala met leu val
121/41   151/51
gcc ggc gtg ttc ggt cac ctg acg gtc ggg atg ttc ttg ggt ctc ggg ttg ctg ctg ggt
ala gly val phe gly his leu thr val gly met phe leu gly leu gly leu leu leu gly
181/61   211/71
ttg ctc aat gcc ctg ctg gtg cgg cgt tcg gcc gag tcg atc acc gcc aaa gag cac ccg
leu leu asn ala leu leu val arg arg ser ala glu ser ile thr ala lys glu his pro
241/81   271/91
tta aaa cgg tcg atg gcc ctc aac tcg gca tcg cga ctg gcg att atc acc atc ctc ggg
leu lys arg ser met ala leu asn ser ala ser arg leu ala ile ile thr ile leu gly
301/101   331/111
ctg atc atc gcc tac att ttc cgg ccc gct gga ttg ggc gtc gtg ttc ggg ctg gca ttc
leu ile ile ala tyr ile phe arg pro ala gly leu gly val val phe gly leu ala phe
361/121   391/131
ttc cag gtg ctg ctg gtg gca acg acg gcc ctg ccg gtc ctg aag aag ctg cgc act gcg
phe gln val leu leu val ala thr thr ala leu pro val leu lys lys leu arg thr ala
421/141   451/151
acc gag gaa ccg gtc gca act tat tct tcc aat ggc cag acc ggg gga tcg gaa gga agg
thr glu glu pro val ala thr tyr ser ser asn gly gln thr gly gly ser glu gly arg
481/161
agc gcc agc gat gac tga
ser ala ser asp asp)OPA

SEQ ID NOS:385-386

FIG. 17E

SEQ ID NO: 387   31/11
```
GTC GAA CAG GTA CGG AAG GCG CCG TCG GTC    GCT CGG TCC GCT GGT ATC TCG TGT TCA GCC
(val glu gln val arg lys ala pro ser val    ala arg ser ala gly ile ser cys ser ala
61/21  ~SEQ ID NO: 388                      91/31
AGC CAG CGG CCG TTA ACG TGG CCG AAC AGG    TCG TCT TGG GGT CGG GCA TCA GCG TCG ATG
ser gln arg pro leu thr trp pro asn arg    ser ser trp gly arg ala ser ala ser met
121/41                                      151/51
TGG CTC AGG TCG ATA CCC GAG GGG ATG GCA    AGT GTC ACC CCG CCA TCC TTC CAC CTC TTT
trp leu arg ser ile pro glu gly met ala    ser val thr pro pro ser phe his leu phe
181/61                                      211/71
TCG GGT GCA ACG ATC GGG CCA TGC CTG ACG    GGG AGC AGA GCC AGC CAC CGG CCC AAG AAG
ser gly ala thr ile gly pro cys leu thr    gly ser arg ala ser his arg pro lys lys
241/81                                      271/91
ATG CGG AAG ACG ACT CGC GGC CCG ACG CCG    CGG AGG CCG CCG CGG CCG AAC CCA AAT CAT
met arg lys thr thr arg gly pro thr pro    arg arg pro pro arg pro asn pro asn his
301/101                                     331/111
CAG CCG GTC CCG ATG TTC TCG ACC TAC GGT    ATC GCC TCG ACA CTA CTC GGC GTG CTA TCG
gln pro val pro met phe ser thr tyr gly    ile ala ser thr leu leu gly val leu ser
361/121
GTC GCC GCG GTC GTG CTG GGT GCG ATG ATC
val ala ala val val leu gly ala met ile)
```

SEQ ID NOS:387-388

FIG. 18A

SEQ ID NO: 389   32/11
```
TCG AAC AGG TAC GGA AGG CGC CGT CGG TCG    CTC GGT CCG CTG GTA TCT CGT GTT CAG CCA
(ser asn arg tyr gly arg arg arg arg ser    leu gly pro leu val ser arg val gln pro
62/21  ~SEQ ID NO: 390                      92/31
GCC AGC GGC CGT TAA CGT GGC CGA ACA GGT    CGT CTT GGG GTC GGG CAT CAG CGT CGA TGT
ala ser gly arg)OCH(arg gly arg thr gly    arg leu gly val gly his gln arg arg cys
122/41       ~SEQ ID NO: 391  152/51
GGC TCA GGT CGA TAC CCG AGG GGA TGG CAA    GTG TCA CCC CGC CAT CCT TCC ACC TCT TTT
gly ser gly arg tyr pro arg gly trp gln    val ser pro arg his pro ser thr ser phe
182/61                                      212/71
CGG GTG CAA CGA TCG GGC CAT GCC TGA CGG    GGA GCA GAG CCA GCC ACC GGC CCA AGA AGA
arg val gln arg ser gly his ala)OPA(arg    gly ala glu pro ala thr gly pro arg arg
242/81       SEQ ID NO: 392                 272/91
TGC GGA AGA CGA CTC GCG GCC CGA CGC CGC    GGA GGC CGC CGC GGC CGA ACC CAA ATC ATC
cys gly arg arg leu ala ala arg arg arg    gly gly arg arg gly arg thr gln ile ile
302/101                                     332/111
AGC CGG TCC CGA TGT TCT CGA CCT ACG GTA    TCG CCT CGA CAC TAC TCG GCG TGC TAT CGG
ser arg ser arg cys ser arg pro thr val    ser pro arg his tyr ser ala cys tyr arg
362/121
TCG CCG CGG TCG TGC TGG GTG CGA TGA TC
ser pro arg ser cys trp val arg)OPA
```

SEQ ID NOS:389-392

FIG. 18B

3/1 ___ SEQ ID NO: 393   33/11
CGA ACA GGT ACG GAA GGC GCC GTC GGT CGC   TCG GTC CGC TGG TAT CTC GTG TTC AGC CAG
(arg thr gly thr glu gly ala val gly arg   ser val arg trp tyr leu val phe ser gln
63/21 ___ SEQ ID NO: 394   93/31
CCA GCG GCC GTT AAC GTG GCC GAA CAG GTC   GTC TTG GGG TCG GGC ATC AGC GTC GAT GTG
pro ala ala val asn val ala glu gln val   val leu gly ser gly ile ser val asp val
123/41   153/51
GCT CAG GTC GAT ACC CGA GGG GAT GGC AAG   TGT CAC CCC GCC ATC CTT CCA CCT CTT TTC
ala gln val asp thr arg gly asp gly lys   cys his pro ala ile leu pro pro leu phe
183/61   213/71
GGG TGC AAC GAT CGG GCC ATG CCT GAC GGG   GAG CAG AGC CAG CCA CCG GCC CAA GAA GAT
gly cys asn asp arg ala met pro asp gly   glu gln ser gln pro pro ala gln glu asp
243/81   273/91
GCG GAA GAC GAC TCG CGG CCC GAC GCC GCG   GAG GCC GCC GCG GCC GAA CCC AAA TCA TCA
ala glu asp asp ser arg pro asp ala ala   glu ala ala ala ala glu pro lys ser ser
303/101   333/111
GCC GGT CCC GAT GTT CTC GAC CTA CGG TAT   CGC CTC GAC ACT ACT CGG CGT GCT ATC GGT
ala gly pro asp val leu asp leu arg tyr   arg leu asp thr thr arg arg ala ile gly
363/121
CGC CGC GGT CGT GCT GGG TGC GAT GAT C
arg arg gly arg ala gly cys asp asp)

SEQ ID NOS:393-394

FIG. 18C

part of the nucleotide sequence of seq18A
1/1 ___ SEQ ID NO: 395   31/11
GAA GGC GCC GTC GGT CGC TCG GTC CGC TGG   TAT CTC GTG TTC AGC CAG CCA GCG GCC GTT
(glu gly ala val gly arg ser val arg trp   tyr leu val phe ser gln pro ala ala val
61/21 ___ SEQ ID NO: 396   91/31
AAC GTG GCC GAA CAG GTC GTC TTG GGG TCG   GGC ATC AGC GTC GAT GTG GCT CAG GTC GAT
asn val ala glu gln val val leu gly ser   gly ile ser val asp val ala gln val asp
121/41   151/51
ACC CGA GGG GAT GGC AAG TGT CAC CCC GCC   ATC CTT CCA CCT CTT TTC GGG TGC AAC GAT
thr arg gly asp gly lys cys his pro ala   ile leu pro pro leu phe gly cys asn asp
181/61   211/71
CGG GCC ATG CCT GAC GGG GAG CAG AGC CAG   CCA CCG GCC CAA GAA GAT GCG GAA GAC GAC
arg ala met pro asp gly glu gln ser gln   pro pro ala gln glu asp ala glu asp asp
241/81   271/91
TCG CGG CCC GAC GCC GCG GAG GCC GCC GCG   GCC GAA CCC AAA TCA TCA GCC GGT CCG ATG
ser arg pro asp ala ala glu ala ala ala   ala glu pro lys ser ser ala gly pro met
301/101   331/111
TTC TCG ACC TAC GGT ATC GCC TCG ACA CTA   CTC GGC GTG CTA TCG GTC GCC GCG GTC GTG
phe ser thr tyr gly ile ala ser thr leu   leu gly val leu ser val ala ala val val
361/121
CTG GGT GCG ATG ATC
leu gly ala met ile)

SEQ ID NOS:395-396

FIG. 18A'

1/1 ___—SEQ ID NO: 397   31/11
CGG AAG GCG CCG TCG GTC GCT CGG TCC GCT GGT ATC TCG TGT TCA GCC AGC CAG CGG CCG
(arg lys ala pro ser val ala arg ser ala gly ile ser cys ser ala ser gln arg pro
61/21 ___—SEQID NO: 398   91/31
TTA ACG TGG CCG AAC AGG TCG TCT TGG GGT CGG GCA TCA GCG TCG ATG TGG CTC AGG TCG
leu thr trp pro asn arg ser ser trp gly arg ala ser ala ser met trp leu arg ser
121/41   151/51
ATA CCC GAG GGG ATG GCA AGT GTC ACC CCG CCA TCC TTC CAC CTC TTT TCG GGT GCA ACG
ile pro glu gly met ala ser val thr pro pro ser phe his leu phe ser gly ala thr
181/61   211/71
ATC GGG CCA TGC CTG ACG GGG AGC AGA GCC AGC CAC CGG CCC AAG AAG ATG CGG AAG ACG
ile gly pro cys leu thr gly ser arg ala ser his arg pro lys lys met arg lys thr
241/81   271/91
ACT CGC GGC CCG ACG CCG CGG AGG CCG CCG CGG CCG AAC CCA AAT CAT CAG CCG GTC CGA
thr arg gly pro thr pro arg arg pro pro arg pro asn pro asn his gln pro val arg
301/101   331/111
TGT TCT CGA CCT ACG GTA TCG CCT CGA CAC TAC TCG GCG TGC TAT CGG TCG CCG CGG TCG
cys ser arg pro thr val ser pro arg his tyr ser ala cys tyr arg ser pro arg ser
361/121
TGC TGG GTG CGA TGA TC
cys trp val arg)OPA

SEQ ID NOS: 397-398

FIG. 18B'

1/1 ___—SEQ ID NO: 399   31/11
GGA AGG CGC CGT CGG TCG CTC GGT CCG CTG GTA TCT CGT GTT CAG CCA GCC AGC GGC CGT
(gly arg arg arg arg ser leu gly pro leu val ser arg val gln pro ala ser gly arg)
61/21 ___—SEQ ID NO: 400   91/31
TAA CGT GGC CGA ACA GGT CGT CTT GGG GTC GGG CAT CAG CGT CGA TGT GGC TCA GGT CGA
OCH(arg gly arg thr gly arg leu gly val gly his gln arg arg cys gly ser gly arg
121/41 ___—SEQ ID NO: 401   151/51
TAC CCG AGG GGA TGG CAA GTG TCA CCC CGC CAT CCT TCC ACC TCT TTT CGG GTG CAA CGA
tyr pro arg gly trp gln val ser pro arg his pro ser thr ser phe arg val gln arg
181/61   211/71
TCG GGC CAT GCC TGA CGG GGA GCA GAG CCA GCC ACC GGC CCA AGA AGA TGC GGA AGA CGA
ser gly his ala)OPA(arg gly ala glu pro ala thr gly pro arg arg cys gly arg arg
241/81   ___—SEQ ID NO: 402   271/91
CTC GCG GCC CGA CGC CGC GGA GGC CGC CGC GGC CGA ACC CAA ATC ATC AGC CGG TCC GAT
leu ala ala arg arg arg gly gly arg arg gly arg thr gln ile ile ser arg ser asp
301/101   331/111
GTT CTC GAC CTA CGG TAT CGC CTC GAC ACT ACT CGG CGT GCT ATC GGT CGC CGC GGT CGT
val leu asp leu arg tyr arg leu asp thr thr arg arg ala ile gly arg arg gly arg
361/121
GCT GGG TGC GAT GAT C
ala gly cys asp asp)

SEQ ID NOS:399-402

FIG. 18C'

sequence Rv0199 predicted by Cole et al. (Nature 393:537-544) and
containing seq18A'

1/1 SEQ ID NO: 403   31/11
atg cct gac ggg gag cag agc cag cca ccg gcc caa gaa gat gcg gaa gac gac tcg cgg
(Met pro asp gly glu gln ser gln pro pro ala gln glu asp ala glu asp asp ser arg

61/21 SEQ ID NO: 404   91/31
ccc gac gcc gcg gag gcc gcc gcg gcc gaa ccc aaa tca tca gcc ggt ccg atg ttc tcg
pro asp ala ala glu ala ala ala ala glu pro lys ser ser ala gly pro met phe ser

121/41   151/51
acc tac ggt atc gcc tcg aca cta ctc ggc gtg cta tcg gtc gcc gcg gtc gtg ctg ggt
thr tyr gly ile ala ser thr leu leu gly val leu ser val ala ala val val leu gly

181/61   211/71
gcg atg atc tgg tcc gca cac cgc gat gac tcc ggc gag cgt acc tac ctg acc cgg gtc
ala met ile trp ser ala his arg asp asp ser gly glu arg thr tyr leu thr arg val

241/81   271/91
atg ctg acc gcc gct gaa tgg acg gcc gtg ctg atc aac atg aac gcc gac aac atc gat
met leu thr ala ala glu trp thr ala val leu ile asn met asn ala asp asn ile asp

301/101   331/111
gcc agc ctg cag cga ctg cac gac gga acg gtc ggt caa ctc aac acc gac ttc gac gct
ala ser leu gln arg leu his asp gly thr val gly gln leu asn thr asp phe asp ala

361/121   391/131
gtc gtg cag ccc tac cgg cag gtg gtg gag aag ttg cgg acg cac agc agc ggc agg atc
val val gln pro tyr arg gln val val glu lys leu arg thr his ser ser gly arg ile

421/141   451/151
gag gcg gta gcg atc gat acg gtg cac cgc gag ctg gat acc cag tcc ggt gcc gcc cga
glu ala val ala ile asp thr val his arg glu leu asp thr gln ser gly ala ala arg

481/161   511/171
ccg gta gta acc acg aaa ttg cca ccg ttt gcc act cgc acc gac tcg gtg ctg ctg gtc
pro val val thr thr lys leu pro pro phe ala thr arg thr asp ser val leu leu val

541/181   571/191
gcg acg tcg gtc agt gag aac gcc ggc gcc aaa ccc cag acc gtg cac tgg aac ttg cgg
ala thr ser val ser glu asn ala gly ala lys pro gln thr val his trp asn leu arg

601/201   631/211
ctc gat gtc tcc gat gtg gac ggc aag ctg atg atc tcc cgg ttg gag tcg att cga tga
leu asp val ser asp val asp gly lys leu met ile ser arg leu glu ser ile arg)OPA

SEQ ID NOS:403-404

FIG. 18D

ORF according to Cole et al. (Nature 393:537-544) and containing Rv0199

```
1/1  SEQ ID NO: 405                          31/11
taa tcc gat gcc gga ttg ggt gaa atg cac caa gta acg ggt cga gtc ttt gga atc ggt
OCH(ser asp ala gly leu gly glu met his gln val thr gly arg val phe gly ile gly
61/21  SEQ ID NO: 406                        91/31
atc gac ata gac tcc gat gcc gcc gcc cac gcc ggc acg ttg cag agt gcc aag ggc ggc
ile asp ile asp ser asp ala ala ala his ala gly thr leu gln ser ala lys gly gly
121/41                                       151/51
ggc caa ttc ggt ggc gtc ggc cgc gct gtc aat cgt ggc caa ttc gtc gtg cag cgg ttg
gly gln phe gly gly val gly arg ala val asn arg gly gln phe val val gln arg leu
181/61                                       211/71
cac ccc tgc gcg ctc gac ggc ttc ctc gtc gag gaa gct ggc gta gag gtc gcc gat gcg
his pro cys ala leu asp gly phe leu val glu glu ala gly val glu val ala asp ala
241/81                                       271/91
ctg cgc atc ggt gcc tac cgc agc acc tgc ttg gct ggc ctg gat gat cag gtc tcg cac
leu arg ile gly ala tyr arg ser thr cys leu ala gly leu asp asp gln val ser his
301/101                                      331/111
ttg tgt ctc ggc gcg gtc gaa cag gct acg gaa ggc gcc gtc ggt cgc tcg gtc cgc tgg
leu cys leu gly ala val glu gln ala thr glu gly ala val gly arg ser val arg trp
361/121                                      391/131
tat ctc gtg ttc agc cag cca gcg gcc gtt aac gtg gcc gaa cag gtc gtc ttg ggg tcg
tyr leu val phe ser gln pro ala ala val asn val ala glu gln val val leu gly ser
421/141                                      451/151
ggc atc agc gtc gat gtg gct cag gtc gat acc cga ggg gat ggc aag tgt cac ccc gcc
gly ile ser val asp val ala gln val asp thr arg gly asp gly lys cys his pro ala
481/161                                      511/171
atc ctt cca cct ctt ttc ggg tgc aac gat cgg gcc atg cct gac ggg gag cag agc cag
ile leu pro pro leu phe gly cys asn asp arg ala met pro asp gly glu gln ser gln
541/181                                      571/191
cca ccg gcc caa gaa gat gcg gaa gac gac tcg cgg ccc gac gcc gcg gag gcc gcc gcg
pro pro ala gln glu asp ala glu asp asp ser arg pro asp ala ala glu ala ala ala
601/201                                      631/211
gcc gaa ccc aaa tca tca gcc ggt ccg atg ttc tcg acc tac ggt atc gcc tcg aca cta
ala glu pro lys ser ser ala gly pro met phe ser thr tyr gly ile ala ser thr leu
661/221                                      691/231
ctc ggc gtg cta tcg gtc gcc gcg gtc gtg ctg ggt gcg atg atc tgg tcc gca cac cgc
leu gly val leu ser val ala ala val val leu gly ala met ile trp ser ala his arg
721/241                                      751/251
gat gac tcc ggc gag cgt acc tac ctg acc cgg gtc atg ctg acc gcc gct gaa tgg acg
asp asp ser gly glu arg thr tyr leu thr arg val met leu thr ala ala glu trp thr
781/261                                      811/271
gcc gtg ctg atc aac atg aac gcc gac aac atc gat gcc agc ctg cag cga ctg cac gac
ala val leu ile asn met asn ala asp asn ile asp ala ser leu gln arg leu his asp
841/281                                      871/291
gga acg gtc ggt caa ctc aac acc gac ttc gac gct gtc gtg cag ccc tac cgg cag gtg
gly thr val gly gln leu asn thr asp phe asp ala val val gln pro tyr arg gln val
901/301                                      931/311
gtg gag aag ttg cgg acg cac agc agc ggc agg atc gag gcg gta gcg atc gat acg gtg
val glu lys leu arg thr his ser ser gly arg ile glu ala val ala ile asp thr val
961/321                                      991/331
cac cgc gag ctg gat acc cag tcc ggt gcc gcc cga ccg gta gta acc acg aaa ttg cca
his arg glu leu asp thr gln ser gly ala ala arg pro val val thr thr lys leu pro
1021/341                                     1051/351
ccg ttt gcc act cgc acc gac tcg gtg ctg ctg gtc gcg acg tcg gtc agt gag aac gcc
pro phe ala thr arg thr asp ser val leu leu val ala thr ser val ser glu asn ala
1081/361                                     1111/371
ggc gcc aaa ccc cag acc gtg cac tgg aac ttg cgg ctc gat gtc tcc gat gtg gac ggc
gly ala lys pro gln thr val his trp asn leu arg leu asp val ser asp val asp gly
1141/381                                     1171/391
aag ctg atg atc tcc cgg ttg gag tcg att cga tga
lys leu met ile ser arg leu glu ser ile arg)OPA
```

SEQ ID NOS:405-406

FIG. 18E

SEQ ID NO: 407   31/11
GTT GCG CAA CGG GGT GAG CAC CGA CGC GAT GAT GGC GCA ACT ATC GAA ACT GCA GGA CAT
(val ala gln arg gly glu his arg arg asp asp gly ala thr ile glu thr ala gly his
61/21   SEQ ID NO: 408   91/31
CGC CAA CGC CAA CGA CGG CAC TCG CGC GGT GGG CAC CCC TGG CTA TCA GGC CAG CGT CGA
arg gln arg gln arg arg his ser arg gly gly his pro trp leu ser gly gln arg arg
121/41   151/51
CTA TGT GGT AAA CAC ACT GCG CAA CAG CGG TTT TGA TGT GCA AAC CCC GGA GTT CTC CGC
leu cys gly lys his thr ala gln gln arg phe)OPA(cys ala asn pro gly val leu arg
181/61   211/71   SEQ ID NO: 409
TCG CGT GTT CAA GGC CGA AAA AGG GGT GGT GAC CCT CGG CGG CAA CAC CGT GGA GGC GAG
ser arg val gln gly arg lys arg gly gly asp pro arg arg gln his arg gly gly glu
241/81   271/91
GGC GCT CGA GTA CAG CCT CGG CAC ACC GCC GGA CGG GGT GAC GGG CCC GCT GGT GGC TGC
gly ala arg val gln pro arg his thr ala gly arg gly asp gly pro ala gly gly cys
301/101   331/111
CCC CGC CGA CGA CAG TCC GGG CTG CAG TCC GTC GGA CTA CGA CAG GCT GCC GGT GTC CGG
pro arg arg arg gln ser gly leu gln ser val gly leu arg gln ala ala gly val arg
361/121
TGC GGT GGT GCT GGT AGA TC
cys gly gly ala gly arg)

SEQ ID NOS: 407-409

FIG. 19A

SEQ ID NO: 410   32/11
TTG CGC AAC GGG GTG AGC ACC GAC GCG ATG ATG GCG CAA CTA TCG AAA CTG CAG GAC ATC
(leu arg asn gly val ser thr asp ala met met ala gln leu ser lys leu gln asp ile
62/21   SEQ ID NO: 411   92/31
GCC AAC GCC AAC GAC GGC ACT CGC GCG GTG GGC ACC CCT GGC TAT CAG GCC AGC GTC GAC
ala asn ala asn asp gly thr arg ala val gly thr pro gly tyr gln ala ser val asp
122/41   152/51
TAT GTG GTA AAC ACA CTG CGC AAC AGC GGT TTT GAT GTG CAA ACC CCG GAG TTC TCC GCT
tyr val val asn thr leu arg asn ser gly phe asp val gln thr pro glu phe ser ala
182/61   212/71
CGC GTG TTC AAG GCC GAA AAA GGG GTG GTG ACC CTC GGC GGC AAC ACC GTG GAG GCG AGG
arg val phe lys ala glu lys gly val val thr leu gly gly asn thr val glu ala arg
242/81   272/91
GCG CTC GAG TAC AGC CTC GGC ACA CCG CCG GAC GGG GTG ACG GGC CCG CTG GTG GCT GCC
ala leu glu tyr ser leu gly thr pro pro asp gly val thr gly pro leu val ala ala
302/101   332/111
CCC GCC GAC GAC AGT CCG GGC TGC AGT CCG TCG GAC TAC GAC AGG CTG CCG GTG TCC GGT
pro ala asp asp ser pro gly cys ser pro ser asp tyr asp arg leu pro val ser gly
362/121
GCG GTG GTG CTG GTA GAT C
ala val val leu val asp)

SEQ ID NOS:410-411

FIG. 19B

SEQ ID NO: 412   SEQ ID NO: 414   33/11
TGC GCA ACG GGG TGA GCA CCG ACG CGA TGA TGG CGC AAC TAT CGA AAC TGC AGG ACA TCG
(cys ala thr gly)OPA(ala pro thr arg)OPA(trp arg asn tyr arg asn cys arg thr ser
63/21   SEQ ID NO: 413   SEQ ID NO: 415 93/31
CCA ACG CCA ACG ACG GCA CTC GCG CGG TGG GCA CCC CTG GCT ATC AGG CCA GCG TCG ACT
pro thr pro thr thr ala leu ala arg trp ala pro leu ala ile arg pro ala ser thr
123/41   153/51
ATG TGG TAA ACA CAC TGC GCA ACA GCG GTT TTG ATG TGC AAA CCC CGG AGT TCT CCG CTC
met trp)OCH(thr his cys ala thr ala val leu met cys lys pro arg ser ser pro leu
183/61   SEQ ID NO: 416   213/71
GCG TGT TCA AGG CCG AAA AAG GGG TGG TGA CCC TCG GCG GCA ACA CCG TGG AGG CGA GGG
ala cys ser arg pro lys lys gly trp)OPA(pro ser ala ala thr pro trp arg arg gly
243/81   SEQ ID NO: 417   273/91
CGC TCG AGT ACA GCC TCG GCA CAC CGC CGG ACG GGG TGA CGG GCC CGC TGG TGG CTG CCC
arg ser ser thr ala ser ala his arg arg thr gly)OPA(arg ala arg trp trp leu pro
303/101   333/111   SEQ ID NO: 418
CCG CCG ACG ACA GTC CGG GCT GCA GTC CGT CGG ACT ACG ACA GGC TGC CGG TGT CCG GTG
pro pro thr thr val arg ala ala val arg arg thr thr thr gly cys arg cys pro val
363/121
CGG TGG TGC TGG TAG ATC
arg trp cys trp)AMB ile

SEQ ID NOS:412-418

FIG. 19C

part of the nucleotide sequence of seq19A

1/1   SEQ ID NO: 419   31/11
CTA TCG AAA CTG CAG GAC ATC GCC AAC GCC AAC GAC GGC ACT CGC GCG GTG GGC ACC CCT
(leu ser lys leu gln asp ile ala asn ala asn asp gly thr arg ala val gly thr pro
61/21   SEQ ID NO: 420   91/31
GGC TAT CAG GCC AGC GTC GAC TAT GTG GTA AAC ACA CTG CGC AAC AGC GGT TTT GAT GTG
gly tyr gln ala ser val asp tyr val val asn thr leu arg asn ser gly phe asp val
121/41   151/51
CAA ACC CCG GAG TTC TCC GCT CGC GTG TTC AAG GCC GAA AAA GGG GTG GTG ACC CTC GGC
gln thr pro glu phe ser ala arg val phe lys ala glu lys gly val val thr leu gly
181/61   211/71
GGC AAC ACC GTG GAG GCG AGG GCG CTC GAG TAC AGC CTC GGC ACA CCG CCG GAC GGG GTG
gly asn thr val glu ala arg ala leu glu tyr ser leu gly thr pro pro asp gly val
241/81   271/91
ACG GGC CCG CTG GTG GCT GCC CCC GCC GAC GAC AGT CCG GGC TGC AGT CCG TCG GAC TAC
thr gly pro leu val ala ala pro ala asp asp ser pro gly cys ser pro ser asp tyr
301/101   331/111
GAC AGG CTG CCG GTG TCC GGT GCG GTG GTG CTG GTA GAT C
asp arg leu pro val ser gly ala val val leu val asp)

SEQ ID No419-420

FIG. 19A'

1/1 ⟋SEQ ID NO: 421   31/11
TAT CGA AAC TGC AGG ACA TCG CCA ACG CCA ACG ACG GCA CTC GCG CGG TGG GCA CCC CTG
(tyr arg asn cys arg thr ser pro thr pro thr thr ala leu ala arg trp ala pro leu
61/21 ⟋SEQ ID NO: 422   91/31
GCT ATC AGG CCA GCG TCG ACT ATG TGG TAA ACA CAC TGC GCA ACA GCG GTT TTG ATG TGC
ala ile arg pro ala ser thr met trp)OCH(thr his cys ala thr ala val leu met cys
121/41   SEQ ID NO: 423 ⟋151/51
AAA CCC CGG AGT TCT CCG CTC GCG TGT TCA AGG CCG AAA AAG GGG TGG TGA CCC TCG GCG
lys pro arg ser ser pro leu ala cys ser arg pro lys lys gly trp)OPA(pro ser ala
181/61   211/71   SEQ ID NO: 424 ⟋
GCA ACA CCG TGG AGG CGA GGG CGC TCG AGT ACA GCC TCG GCA CAC CGC CGG ACG GGG TGA
ala thr pro trp arg arg gly arg ser ser thr ala ser ala his arg arg thr gly)OPA
241/81   271/91
CGG GCC CGC TGG TGG CTG CCC CCG CCG ACG ACA GTC CGG GCT GCA GTC CGT CGG ACT ACG
(arg ala arg trp trp leu pro pro pro thr thr val arg ala ala val arg arg thr thr
301/101 ⟋SEQ ID NO: 425   331/111
ACA GGC TGC CGG TGT CCG GTG CGG TGG TGC TGG TAG ATC
thr gly cys arg cys pro val arg trp cys trp)AMB ile

SEQ ID NOS:421-425

FIG. 19B'

1/1 ⟋SEQ ID NO: 426   31/11
ATC GAA ACT GCA GGA CAT CGC CAA CGC CAA CGA CGG CAC TCG CGC GGT GGG CAC CCC TGG
(ile glu thr ala gly his arg gln arg gln arg arg his ser arg gly gly his pro trp
61/21 ⟋SEQ ID NO: 427   91/31
CTA TCA GGC CAG CGT CGA CTA TGT GGT AAA CAC ACT GCG CAA CAG CGG TTT TGA TGT GCA
leu ser gly gln arg arg leu cys gly lys his thr ala gln gln arg phe)OPA(cys ala
121/41   151/51   SEQ ID NO: 428 ⟋
AAC CCC GGA GTT CTC CGC TCG CGT GTT CAA GGC CGA AAA AGG GGT GGT GAC CCT CGG CGG
asn pro gly val leu arg ser arg val gln gly arg lys arg gly gly asp pro arg arg
181/61   211/71
CAA CAC CGT GGA GGC GAG GGC GCT CGA GTA CAG CCT CGG CAC ACC GCC GGA CGG GGT GAC
gln his arg gly gly glu gly ala arg val gln pro arg his thr ala gly arg gly asp
241/81   271/91
GGG CCC GCT GGT GGC TGC CCC CGC CGA CGA CAG TCC GGG CTG CAG TCC GTC GGA CTA CGA
gly pro ala gly gly cys pro arg arg arg gln ser gly leu gln ser val gly leu arg
301/101   331/111
CAG GCT GCC GGT GTC CGG TGC GGT GGT GCT GGT AGA TC
gln ala ala gly val arg cys gly gly ala gly arg)

SEQ ID NOS:426-428

FIG. 19C'

sequence Rv0418 predicted by Cole et al. (Nature 393:537-544) and
containing seq19A'

1/1 ___— SEQ ID NO: 429   31/11
atg gtg aac aaa tcc agg atg atg ccg gcg gtg ctg gcc gtg gct gtg gtc gtc gca ttc
(Met val asn lys ser arg met met pro ala val leu ala val ala val val val gca phe
61/21 ___— SEQ ID NO: 430   91/31
ctg acg acg ggc tgt atc cgg tgg tct acg cag tcg cgg ccc gtt gtt aac ggc ccc gct
leu thr thr gly cys ile arg trp ser thr gln ser arg pro val val asn gly pro ala
121/41   151/51
gcc gca gag ttc gcc gtt gcg ttg cgc aac cgg gtg agc acc gac gcg atg atg gcg cac
ala ala glu phe ala val ala leu arg asn arg val ser thr asp ala met met ala his
181/61   211/71
cta tcg aaa ctg cag gac atc gcc aac gcc aac gac ggc act cgc gcg gtg ggc acc cct
leu ser lys leu gln asp ile ala asn ala asn asp gly thr arg ala val gly thr pro
241/81   271/91
ggc tat cag gcc agc gtc gac tat gtg gta aac aca ctg cgc aac agc ggt ttt gat gtg
gly tyr gln ala ser val asp tyr val val asn thr leu arg asn ser gly phe asp val
301/101   331/111
caa acc ccg gag ttc tcc gct cgc gtg ttc aag gcc gaa aaa ggg gtg gtg acc ctc ggc
gln thr pro glu phe ser ala arg val phe lys ala glu lys gly val val thr leu gly
361/121   391/131
ggc aac acc gtg gag gcg agg gcg ctc gag tac agc ctc ggc aca ccg ccg gac ggg gtg
gly asn thr val glu ala arg ala leu glu tyr ser leu gly thr pro pro asp gly val
421/141   451/151
acg ggc ccg ctg gtg gct gcc ccc gcc gac gac agt ccg ggc tgc agt ccg tcg gac tac
thr gly pro leu val ala ala pro ala asp asp ser pro gly cys ser pro ser asp tyr
481/161   511/171
gac agg ctg ccg gtg tcc ggt gcg gtg gtg ctg gta gat cgc ggc gtc tgt cct ttt gcc
asp arg leu pro val ser gly ala val val leu val asp arg gly val cys pro phe ala
541/181   571/191
cag aag gaa gac gca gcc gcg cag cgc ggt gcg gtg gcg ctg atc att gct gac aac atc
gln lys glu asp ala ala ala gln arg gly ala val ala leu ile ile ala asp asn ile
601/201   631/211
gac gag cag gcg atg ggc ggc acc ctg ggg gct aat acc gac gtc aag atc ccg gtg gtg
asp glu gln ala met gly gly thr leu gly ala asn thr asp val lys ile pro val val
661/221   691/231
agt gtc acc aag tcg gtc gga ttc cag cta cgc gga cag tct ggg cca acc acc gtc aag
ser val thr lys ser val gly phe gln leu arg gly gln ser gly pro thr thr val lys
721/241   751/251
ctc acg gcg agc acc caa agt ttc aag gcc cgc aac gtc atc gcg cag acg aag acg ggg
leu thr ala ser thr gln ser phe lys ala arg asn val ile ala gln thr lys thr gly
781/261   811/271
tcg tcg gcc aac gtg gtg atg gca ggt gcg cat ttg gac agc gtt ccg gaa gga ccc ggc
ser ser ala asn val val met ala gly ala his leu asp ser val pro glu gly pro gly
841/281   871/291
atc aac gac aac ggc tcg gga gtg gct gcg gtt ctg gaa acg gca gtg cag ctg ggg aac
ile asn asp asn gly ser gly val ala ala val leu glu thr ala val gln leu gly asn
901/301   931/311
tca ccg cat gtg tcc aac gcg gta cgg ttc gcc ttc tgg ggc gcc gag gaa ttc ggc ctg
ser pro his val ser asn ala val arg phe ala phe trp gly ala glu glu phe gly leu
961/321   991/331
att ggg tca cga aac tac gtc gag tcg ctg gac atc gac gcg ctc aaa ggc atc gcg ctg
ile gly ser arg asn tyr val glu ser leu asp ile asp ala leu lys gly ile ala leu

SEQ ID NOS:429-430

FIG. 19D

1021/341
tat ctg aac ttc gac atg ttg gcg tcg ccg
tyr leu asn phe asp met leu ala ser pro
1081/361
cag tcg ctg ccg cta gac gcc cgc ggt cag
gln ser leu pro leu asp ala arg gly gln
1141/381
gag cgc acg ttc gtc gcc tat ctg aag atg
glu arg thr phe val ala tyr leu lys met
1201/401
gac ggt cgg tcc gac tac gac ggc ttc acg
asp gly arg ser asp tyr asp gly phe thr
1261/421
tcc ggc gct gag gtc aag aag tcc gcc gag
ser gly ala glu val lys lys ser ala glu
1321/441
gag cct ttc gat ccc aac tat cac cag aag
glu pro phe asp pro asn tyr his gln lys
1381/461
gcg ctc ggt atc aac ggc gct ggc gtc gcg
ala leu gly ile asn gly ala gly val ala
1441/481
ggc ggc ccc aac ggg gtt ccg gtc atg gcg
gly gly pro asn gly val pro val met ala
1501/501
tga
OPA

1051/351
aac ccg ggt tac ttc acc tac gac ggt gac
asn pro gly tyr phe thr tyr asp gly asp
1111/371
ccg gtg gtg ccc gaa ggc tcg gcc ggt atc
pro val val pro glu gly ser ala gly ile
1171/391
gcc ggc aag acc gcg cag gac acc tcg ttc
ala gly lys thr ala gln asp thr ser phe
1231/411
ctg gcg ggt atc cct tcg ggt ggc ctg ttc
leu ala gly ile pro ser gly gly leu phe
1291/431
caa gcc gag ctc tgg ggc ggc acc gcc gac
gln ala glu leu trp gly gly thr ala asp
1351/451
aca gac acc ctg gac cat atc gac cgc acc
thr asp thr leu asp his ile asp arg thr
1411/471
tac gcg gtg ggt ttg tat gcg cag gac ctc
tyr ala val gly leu tyr ala gln asp leu
1471/491
gac cgc acc cgc cac ctg att gcc aaa ccg
asp arg thr arg his leu ile ala lys pro)

SEQ ID NOS: 429-430 (continued)

FIG. 19D (continued)

ORF according to Cole et al. (Nature 393:537-544) and containing Rv0418

1/1 SEQ ID NO: 431										31/11									
tag	gcc	att	caa	cgc	tct	gtt	cgt	ttg	att	ggt	cgg	tgg	gat	gcg	aaa	gct	gcg	cgg	cga
AMB(ala	ile	gln	arg	ser	val	arg	leu	ile	gly	arg	trp	asp	ala	lys	ala	ala	arg	arg
61/21 SEQ ID NO: 432										91/31									
cag	gcg	cgg	tct	aat	ctg	ggc	gcg	atg	gtg	aac	aaa	tcc	agg	atg	atg	ccg	gcg	gtg	ctg
gln	ala	arg	ser	asn	leu	gly	ala	met	val	asn	lys	ser	arg	met	met	pro	ala	val	leu
121/41										151/51									
gcc	gtg	gct	gtg	gtc	gtc	gca	ttc	ctg	acg	acg	ggc	tgt	atc	cgg	tgg	tct	acg	cag	tcg
ala	val	ala	val	val	val	ala	phe	leu	thr	thr	gly	cys	ile	arg	trp	ser	thr	gln	ser
181/61										211/71									
cgg	ccc	gtt	gtt	aac	ggc	ccc	gct	gcc	gca	gag	ttc	gcc	gtt	gcg	ttg	cgc	aac	cgg	gtg
arg	pro	val	val	asn	gly	pro	ala	ala	ala	glu	phe	ala	val	ala	leu	arg	asn	arg	val
241/81										271/91									
agc	acc	gac	gcg	atg	atg	gcg	cac	cta	tcg	aaa	ctg	cag	gac	atc	gcc	aac	gcc	aac	gac
ser	thr	asp	ala	met	met	ala	his	leu	ser	lys	leu	gln	asp	ile	ala	asn	ala	asn	asp
301/101										331/111									
ggc	act	cgc	gcg	gtg	ggc	acc	cct	ggc	tat	cag	gcc	agc	gtc	gac	tat	gtg	gta	aac	aca
gly	thr	arg	ala	val	gly	thr	pro	gly	tyr	gln	ala	ser	val	asp	tyr	val	val	asn	thr
361/121										391/131									
ctg	cgc	aac	agc	ggt	ttt	gat	gtg	caa	acc	ccg	gag	ttc	tcc	gct	cgc	gtg	ttc	aag	gcc
leu	arg	asn	ser	gly	phe	asp	val	gln	thr	pro	glu	phe	ser	ala	arg	val	phe	lys	ala
421/141										451/151									
gaa	aaa	ggg	gtg	gtg	acc	ctc	ggc	ggc	aac	acc	gtg	gag	gcg	agg	gcg	ctc	gag	tac	agc
glu	lys	gly	val	val	thr	leu	gly	gly	asn	thr	val	glu	ala	arg	ala	leu	glu	tyr	ser
481/161										511/171									
ctc	ggc	aca	ccg	ccg	gac	ggg	gtg	acg	ggc	ccg	ctg	gtg	gct	gcc	ccc	gcc	gac	gac	agt
leu	gly	thr	pro	pro	asp	gly	val	thr	gly	pro	leu	val	ala	ala	pro	ala	asp	asp	ser
541/181										571/191									
ccg	ggc	tgc	agt	ccg	tcg	gac	tac	gac	agg	ctg	ccg	gtg	tcc	ggt	gcg	gtg	gtg	ctg	gta
pro	gly	cys	ser	pro	ser	asp	tyr	asp	arg	leu	pro	val	ser	gly	ala	val	val	leu	val
601/201										631/211									
gat	cgc	ggc	gtc	tgt	cct	ttt	gcc	cag	aag	gaa	gac	gca	gcc	gcg	cag	cgc	ggt	gcg	gtg
asp	arg	gly	val	cys	pro	phe	ala	gln	lys	glu	asp	ala	ala	ala	gln	arg	gly	ala	val
661/221										691/231									
gcg	ctg	atc	att	gct	gac	aac	atc	gac	gag	cag	gcg	atg	ggc	ggc	acc	ctg	ggg	gct	aat
ala	leu	ile	ile	ala	asp	asn	ile	asp	glu	gln	ala	met	gly	gly	thr	leu	gly	ala	asn
721/241										751/251									
acc	gac	gtc	aag	atc	ccg	gtg	gtg	agt	gtc	acc	aag	tcg	gtc	gga	ttc	cag	cta	cgc	gga
thr	asp	val	lys	ile	pro	val	val	ser	val	thr	lys	ser	val	gly	phe	gln	leu	arg	gly
781/261										811/271									
cag	tct	ggg	cca	acc	acc	gtc	aag	ctc	acg	gcg	agc	acc	caa	agt	ttc	aag	gcc	cgc	aac
gln	ser	gly	pro	thr	thr	val	lys	leu	thr	ala	ser	thr	gln	ser	phe	lys	ala	arg	asn
841/281										871/291									
gtc	atc	gcg	cag	acg	aag	acg	ggg	tcg	tcg	gcc	aac	gtg	gtg	atg	gca	ggt	gcg	cat	ttg
val	ile	ala	gln	thr	lys	thr	gly	ser	ser	ala	asn	val	val	met	ala	gly	ala	his	leu
901/301										931/311									
gac	agc	gtt	ccg	gaa	gga	ccc	ggc	atc	aac	gac	aac	ggc	tcg	gga	gtg	gct	gcg	gtt	ctg
asp	ser	val	pro	glu	gly	pro	gly	ile	asn	asp	asn	gly	ser	gly	val	ala	ala	val	leu
961/321										991/331									
gaa	acg	gca	gtg	cag	ctg	ggg	aac	tca	ccg	cat	gtg	tcc	aac	gcg	gta	cgg	ttc	gcc	ttc
glu	thr	ala	val	gln	leu	gly	asn	ser	pro	his	val	ser	asn	ala	val	arg	phe	ala	phe

SEQ ID NOS:431-432

FIG. 19E

1021/341
tgg ggc gcc gag gaa ttc ggc ctg att ggg
trp gly ala glu glu phe gly leu ile gly
1081/361
gac gcg ctc aaa ggc atc gcg ctg tat ctg
asp ala leu lys gly ile ala leu tyr leu
1141/381
ggt tac ttc acc tac gac ggt gac cag tcg
gly tyr phe thr tyr asp gly asp gln ser
1201/401
gtg ccc gaa ggc tcg gcc ggt atc gag cgc
val pro glu gly ser ala gly ile glu arg
1261/421
aag acc gcg cag gac acc tcg ttc gac ggt
lys thr ala gln asp thr ser phe asp gly
1321/441
ggt atc cct tcg ggt ggc ctg ttc tcc ggc
gly ile pro ser gly gly leu phe ser gly
1381/461
gag ctc tgg ggc ggc acc gcc gac gag cct
glu leu trp gly gly thr ala asp glu pro
1441/481
acc ctg gac cat atc gac cgc acc gcg ctc
thr leu asp his ile asp arg thr ala leu
1501/501
gtg ggt ttg tat gcg cag gac ctc ggc ggc
val gly leu tyr ala gln asp leu gly gly
1561/521
acc cgc cac ctg att gcc aaa ccg tga
thr arg his leu ile ala lys pro)OPA

1051/351
tca cga aac tac gtc gag tcg ctg gac atc
ser arg asn tyr val glu ser leu asp ile
1111/371
aac ttc gac atg ttg gcg tcg ccg aac ccg
asn phe asp met leu ala ser pro asn pro
1171/391
ctg ccg cta gac gcc cgc ggt cag ccg gtg
leu pro leu asp ala arg gly gln pro val
1231/411
acg ttc gtc gcc tat ctg aag atg gcc ggc
thr phe val ala tyr leu lys met ala gly
1291/431
cgg tcc gac tac gac ggc ttc acg ctg gcg
arg ser asp tyr asp gly phe thr leu ala
1351/451
gct gag gtc aag aag tcc gcc gag caa gcc
ala glu val lys lys ser ala glu gln ala
1411/471
ttc gat ccc aac tat cac cag aag aca gac
phe asp pro asn tyr his gln lys thr asp
1471/491
ggt atc aac ggc gct ggc gtc gcg tac gcg
gly ile asn gly ala gly val ala tyr ala
1531/511
ccc aac ggg gtt ccg gtc atg gcg gac cgc
pro asn gly val pro val met ala asp arg

SEQ ID NOS:431-432 (continued)

FIG. 19E (continued)

SEQ ID NO: 433
CGA GAC AGT GGT GCG GGA CAC TTG AGT TCG
(arg asp ser gly ala gly his leu ser ser
61/21 ——SEQ ID NO: 434
CGC GGT GTG GGA CTC ACG TTC GGT GAG GGT
arg gly val gly leu thr phe gly glu gly
121/41
GGC CGG CTG GCA ACC GTC GGT TTC GAC GTT
gly arg leu ala thr val gly phe asp val
181/61
TGA GCT CCC CGT TTT GCT GGA TGC CCA GGC
OPA(ala pro arg phe ala gly cys pro gly
241/81 ——SEQ ID NO: 435
GCA CAT GGT GCC GGC AGG GAG GAA CAG TGG
ala his gly ala gly arg glu glu gln trp
301/101
TCG GTG CGT GCA TGC TCG CAG CCG GAT GCA
ser val arg ala cys ser gln pro asp ala
361/121
CCG ACA AAT CCG GAC CAC TGC ATC AGG ATC
pro thr asn pro asp his cys ile arg ile)

31/11
GCT GCT AAC GAC GCC AGA GTC GCC CGC TTC
ala ala asn asp ala arg val ala arg phe
91/31
ACA GCG GAC CTT CGA GCA CGC AAT ATC GTG
thr ala asp leu arg ala arg asn ile val
151/51
GGT GAC GAC CCC TCG TTC ATG AAT CGT TCT
gly asp asp pro ser phe met asn arg ser)
211/71
ACC GCC GGT ACT GCT GCG CTT AAG CTT GTC
thr ala gly thr ala ala leu lys leu val
271/91
GCA AGC AGC TAG CCG CGC TCG CCG CGC TGG
ala ser ser)AMB(pro arg ser pro arg trp
331/111   ——SEQ ID NO: 436
CCA ACG TGG TCG ACG GGA CCG CCG TGG CTG
pro thr trp ser thr gly pro pro trp leu

SEQ ID NOS:433-436

FIG. 20A

```
 SEQ ID NO: 437      SEQ ID NO: 439      32/11
GAG ACA GTG GTG CGG GAC ACT TGA GTT CGG CTG CTA ACG ACG CCA GAG TCG CCC GCT TCC
(glu thr val val arg asp thr)OPA(val arg leu leu thr thr pro glu ser pro ala ser
62/21   SEQ ID NO: 438                   92/31
GCG GTG TGG GAC TCA CGT TCG GTG AGG GTA CAG CGG ACC TTC GAG CAC GCA ATA TCG TGG
ala val trp asp ser arg ser val arg val gln arg thr phe glu his ala ile ser trp
122/41                                   152/51
GCC GGC TGG CAA CCG TCG GTT TCG ACG TTG GTG ACG ACC CCT CGT TCA TGA ATC GTT CTT
ala gly trp gln pro ser val ser thr leu val thr thr pro arg ser)OPA(ile val leu
182/61                                   212/71     SEQ ID NO: 440
GAG CTC CCC GTT TTG CTG GAT GCC CAG GCA CCG CCG GTA CTG CTG CGC TTA AGC TTG TCG
glu leu pro val leu leu asp ala gln ala pro pro val leu leu arg leu ser leu ser
242/81                                   272/91
CAC ATG GTG CCG GCA GGG AGG AAC AGT GGG CAA GCA GCT AGC CGC GCT CGC CGC GCT GGT
his met val pro ala gly arg asn ser gly gln ala ala ser arg ala arg arg ala gly
302/101                                  332/111
CGG TGC GTG CAT GCT CGC AGC CGG ATG CAC CAA CGT GGT CGA CGG GAC CGC CGT GGC TGC
arg cys val his ala arg ser arg met his gln arg gly arg arg asp arg arg gly cys
362/121
CGA CAA ATC CGG ACC ACT GCA TCA GGA TC
arg gln ile arg thr thr ala ser gly)
```

SEQ ID NOS:437-440

FIG. 20B

```
 SEQ ID NO: 441                          33/11
AGA CAG TGG TGC GGG ACA CTT GAG TTC GGC TGC TAA CGA CGC CAG AGT CGC CCG CTT CCG
(arg gln trp cys gly thr leu glu phe gly cys)OCH(arg arg gln ser arg pro leu pro
63/21   SEQ ID NO: 442                   93/31     SEQ ID NO: 443
CGG TGT GGG ACT CAC GTT CGG TGA GGG TAC AGC GGA CCT TCG AGC ACG CAA TAT CGT GGG
arg cys gly thr his val arg)OPA(gly tyr ser gly pro ser ser thr gln tyr arg gly
123/41            SEQ ID NO: 444         153/51
CCG GCT GGC AAC CGT CGG TTT CGA CGT TGG TGA CGA CCC CTC GTT CAT GAA TCG TTC TTG
pro ala gly asn arg arg phe arg arg trp)OPA(arg pro leu val his glu ser phe leu
183/61                                   213/71     SEQ ID NO: 445
AGC TCC CCG TTT TGC TGG ATG CCC AGG CAC CGC CGG TAC TGC TGC GCT TAA GCT TGT CGC
ser ser pro phe cys trp met pro arg his arg arg tyr cys cys ala)OCH(ala cys arg
243/81                                   273/91     SEQ ID NO: 446
ACA TGG TGC CGG CAG GGA GGA ACA GTG GGC AAG CAG CTA GCC GCG CTC GCC GCG CTG GTC
thr trp cys arg gln gly gly thr val gly lys gln leu ala ala leu ala ala leu val
303/101                                  333/111
GGT GCG TGC ATG CTC GCA GCC GGA TGC ACC AAC GTG GTC GAC GGG ACC GCC GTG GCT GCC
gly ala cys met leu ala ala gly cys thr asn val val asp gly thr ala val ala ala
363/121
GAC AAA TCC GGA CCA CTG CAT CAG GAT C
asp lys ser gly pro leu his gln asp)
```

SEQ ID NOS:441-446

FIG. 20C

part of the nucleotide sequence of seq20A

```
1/1      SEQ ID NO: 447     SEQ ID NO: 449  31/11
TGT GGG ACT CAC GTT CGG TGA GGG TAC AGC GGA CCT TCG AGC ACG CAA TAT CGT GGG CCG
(cys gly thr his val arg)OPA(gly tyr ser gly pro ser ser thr gln tyr arg gly pro
61/21     SEQ ID NO: 448                    91/31
GCT GGC AAC CGT CGG TTT CGA CGT TGG TGA CGA CCC CTC GTT CAT GAA TCG TTC TTG AGC
ala gly asn arg arg phe arg arg trp)OPA(arg pro leu val his glu ser phe leu ser
121/41                 SEQ ID NO: 450       151/51
TCC CCG TTT TGC TGG ATG CCC AGG CAC CGC CGG TAC TGC TGC GCT TAA GCT TGT CGC ACA
ser pro phe cys trp met pro arg his arg arg tyr cys cys ala)OCH(ala cys arg thr
181/61                                      211/71  SEQ ID NO: 451
TGG TGC CGG CAG GGA GGA ACA GTG GGC AAG CAG CTA GCC GCG CTC GCC GCG CTG GTC GGT
trp cys arg gln gly gly thr val gly lys gln leu ala ala leu ala ala leu val gly
241/81                                      271/91
GCG TGC ATG CTC GCA GCC GGA TGC ACC AAC GTG GTC GAC GGG ACC GCC GTG GCT GCC GAC
ala cys met leu ala ala gly cys thr asn val val asp gly thr ala val ala ala asp
301/101
AAA TCC GGA CCA CTG CAT CAG GAT C
lys ser gly pro leu his gln asp)
```

SEQ ID NOS:447-451

FIG. 20A'

```
1/1   SEQ ID NO: 452                       31/11
GTG GGA CTC ACG TTC GGT GAG GGT ACA GCG GAC CTT CGA GCA CGC AAT ATC GTG GGC CGG
(val gly leu thr phe gly glu gly thr ala asp leu arg ala arg asn ile val gly arg
61/21     SEQ ID NO: 453                    91/31
CTG GCA ACC GTC GGT TTC GAC GTT GGT GAC GAC CCC TCG TTC ATG AAT CGT TCT TGA GCT
leu ala thr val gly phe asp val gly asp asp pro ser phe met asn arg ser)OPA(ala
121/41                                      151/51       SEQ ID NO: 454
CCC CGT TTT GCT GGA TGC CCA GGC ACC GCC GGT ACT GCT GCG CTT AAG CTT GTC GCA CAT
pro arg phe ala gly cys pro gly thr ala gly thr ala ala leu lys leu val ala his
181/61                                      211/71
GGT GCC GGC AGG GAG GAA CAG TGG GCA AGC AGC TAG CCG CGC TCG CCG CGC TGG TCG GTG
gly ala gly arg glu glu gln trp ala ser ser)AMB(pro arg ser pro arg trp ser val
241/81                                      271/91     SEQ ID NO: 455
CGT GCA TGC TCG CAG CCG GAT GCA CCA ACG TGG TCG ACG GGA CCG CCG TGG CTG CCG ACA
arg ala cys ser gln pro asp ala pro thr trp ser thr gly pro pro trp leu pro thr
301/101
AAT CCG GAC CAC TGC ATC AGG ATC
asn pro asp his cys ile arg ile)
```

SEQ ID NOS:452-455

FIG. 20B'

```
1/1    ┌─SEQ ID NO: 456                          31/11
GTG TGG GAC TCA CGT TCG GTG AGG GTA CAG  CGG ACC TTC GAG CAC GCA ATA TCG TGG GCC
(val trp asp ser arg ser val arg val gln  arg thr phe glu his ala ile ser trp ala
61/21  ┌─SEQ ID NO: 457                         91/31
GGC TGG CAA CCG TCG GTT TCG ACG TTG GTG  ACG ACC CCT CGT TCA TGA ATC GTT CTT GAG
gly trp gln pro ser val ser thr leu val  thr thr pro arg ser)OPA(ile val leu glu
121/41                                    151/51    SEQ ID NO: 458
CTC CCC GTT TTG CTG GAT GCC CAG GCA CCG  CCG GTA CTG CTG CGC TTA AGC TTG TCG CAC
leu pro val leu leu asp ala gln ala pro  pro val leu leu arg leu ser leu ser his
181/61                                    211/71
ATG GTG CCG GCA GGG AGG AAC AGT GGG CAA  GCA GCT AGC CGC GCT CGC CGC GCT GGT CGG
met val pro ala gly arg asn ser gly gln  ala ala ser arg ala arg arg ala gly arg
241/81                                    271/91
TGC GTG CAT GCT CGC AGC CGG ATG CAC CAA  CGT GGT CGA CGG GAC CGC CGT GGC TGC CGA
cys val his ala arg ser arg met his gln  arg gly arg arg asp arg arg gly cys arg
301/101
CAA ATC CGG ACC ACT GCA TCA GGA TC
gln ile arg thr thr ala ser gly)
```

SEQ ID NOS:456-458

FIG. 20C'

sequence Rv3576 predicted by Cole et al. (Nature 393:537-544) and containing
seq20A'

```
1/1    ┌─SEQ ID NO: 459                          31/11
atg ggc aag cag cta gcc gcg ctc gcc gcg  ctg gtc ggt gcg tgc atg ctc gca gcc gga
(met gly lys gln leu ala ala leu ala ala  leu val gly ala cys met leu ala ala gly
61/21  ┌─SEQ ID NO: 460                         91/31
tgc acc aac gtg gtc gac ggg acc gcc gtg  gct gcc gac aaa tcc gga cca ctg cat cag
cys thr asn val val asp gly thr ala val  ala ala asp lys ser gly pro leu his gln
121/41                                    151/51
gat ccg ata ccg gtt tca gcg ctt gaa ggg  ctg ctt ctc gac ttg agc cag atc aat gcc
asp pro ile pro val ser ala leu glu gly  leu leu leu asp leu ser gln ile asn ala
181/61                                    211/71
gcg ctg ggt gcg aca tcg atg aag gtg tgg  ttc aac gcc aag gca atg tgg gac tgg agc
ala leu gly ala thr ser met lys val trp  phe asn ala lys ala met trp asp trp ser
241/81                                    271/91
aag agc gtg gcc gac aag aat tgc ctg gct  atc gac ggt cca gca cag gaa aag gtc tat
lys ser val ala asp lys asn cys leu ala  ile asp gly pro ala gln glu lys val tyr
301/101                                   331/111
gcc ggc acc ggg tgg acc gct atg cgc ggc  caa cgg ctg gat gac agc atc gat gac tcc
ala gly thr gly trp thr ala met arg gly  gln arg leu asp asp ser ile asp asp ser
361/121                                   391/131
aag aaa cgc gac cac tac gcc att caa gcg  gtc gtc ggc ttc ccg acc gca cat gat gcc
lys lys arg asp his tyr ala ile gln ala  val val gly phe pro thr ala his asp ala
421/141                                   451/151
gag gag ttc tac agc tcc tcg gtg caa agc  tgg agc agc tgc tcg aac cgc cgg ttt gtc
glu glu phe tyr ser ser ser val gln ser  trp ser ser cys ser asn arg arg phe val
481/161                                   511/171
gaa gtc acc ccc gga cag gac gac gcc gcc  tgg act gtg gct gac gtt gtc aac gac aac
glu val thr pro gly gln asp asp ala ala  trp thr val ala asp val val asn asp asn
541/181                                   571/191
ggc atg ctc agt agc tcg cag gtt cag gaa  ggc ggc gac gga tgg acc tgc cag cgt gcc
gly met leu ser ser ser gln val gln glu  gly gly asp gly trp thr cys gln arg ala
601/201                                   631/211
ctg act gcg cgc aac aac gtc act atc gac  att gtc acg tgc gcc tat agc caa ccg gat
leu thr ala arg asn asn val thr ile asp  ile val thr cys ala tyr ser gln pro asp
661/221                                   691/231
ttg gtg gcg att ggc atc gct aac caa atc  gcg gcc aag gtt gct aag cag tag
leu val ala ile gly ile ala asn gln ile  ala ala lys val ala lys gln)AMB
```

SEQ ID NOS:459-460

FIG. 20D

ORF according to Cole et al. (Nature 393:537-544) and containing Rv3576

1/1 —SEQ ID NO: 461   31/11
taa gct tgt cgc aca tgg tgc cgg cag gga gga aca gtg ggc aag cag cta gcc gcg ctc
OCH(ala cys arg thr trp cys arg gln gly gly thr val gly lys gln leu ala ala leu
61/21 —SEQ ID NO: 462   91/31
gcc gcg ctg gtc ggt gcg tgc atg ctc gca gcc gga tgc acc aac gtg gtc gac ggg acc
ala ala leu val gly ala cys met leu ala ala gly cys thr asn val val asp gly thr
121/41   151/51
gcc gtg gct gcc gac aaa tcc gga cca ctg cat cag gat ccg ata ccg gtt tca gcg ctt
ala val ala ala asp lys ser gly pro leu his gln asp pro ile pro val ser ala leu
181/61   211/71
gaa ggg ctg ctt ctc gac ttg agc cag atc aat gcc gcg ctg ggt gcg aca tcg atg aag
glu gly leu leu leu asp leu ser gln ile asn ala ala leu gly ala thr ser met lys
241/81   271/91
gtg tgg ttc aac gcc aag gca atg tgg gac tgg agc aag agc gtg gcc gac aag aat tgc
val trp phe asn ala lys ala met trp asp trp ser lys ser val ala asp lys asn cys
301/101   331/111
ctg gct atc gac ggt cca gca cag gaa aag gtc tat gcc ggc acc ggg tgg acc gct atg
leu ala ile asp gly pro ala gln glu lys val tyr ala gly thr gly trp thr ala met
361/121   391/131
cgc ggc caa cgg ctg gat gac agc atc gat gac tcc aag aaa cgc gac cac tac gcc att
arg gly gln arg leu asp asp ser ile asp asp ser lys lys arg asp his tyr ala ile
421/141   451/151
caa gcg gtc gtc ggc ttc ccg acc gca cat gat gcc gag gag ttc tac agc tcc tcg gtg
gln ala val val gly phe pro thr ala his asp ala glu glu phe tyr ser ser ser val
481/161   511/171
caa agc tgg agc agc tgc tcg aac cgc cgg ttt gtc gaa gtc acc ccc gga cag gac gac
gln ser trp ser ser cys ser asn arg arg phe val glu val thr pro gly gln asp asp
541/181   571/191
gcc gcc tgg act gtg gct gac gtt gtc aac gac aac ggc atg ctc agt agc tcg cag gtt
ala ala trp thr val ala asp val val asn asp asn gly met leu ser ser ser gln val
601/201   631/211
cag gaa ggc ggc gac gga tgg acc tgc cag cgt gcc ctg act gcg cgc aac aac gtc act
gln glu gly gly asp gly trp thr cys gln arg ala leu thr ala arg asn asn val thr
661/221   691/231
atc gac att gtc acg tgc gcc tat agc caa ccg gat ttg gtg gcg att ggc atc gct aac
ile asp ile val thr cys ala tyr ser gln pro asp leu val ala ile gly ile ala asn
721/241
caa atc gcg gcc aag gtt gct aag cag tag
gln ile ala ala lys val ala lys gln)AMB

SEQ ID NOS:461-462

FIG. 20E

1/1 ___SEQ ID NO: 463   31/11
GTC CTG GTC GCC GCG CAA CTG GCC GGT CCC GAT GGA AAG TGT TCA CGA TCG CGC TTC TGC
(val leu val ala ala gln leu ala gly pro asp gly lys cys ser arg ser arg phe cys
61/21 ~SEQ ID NO: 464   SEQ ID NO: 466   91/31
CGC TGG TAG TGG CGA TGG TGT TAG CAG GAT TGC GGG TCG AGG CTG CGA TGG CCA GCA CCA
arg trp)AMB(trp arg trp cys)AMB(gln asp cys gly ser arg leu arg trp pro ala pro
121/41   ~SEQ ID NO: 465   151/51
GCG GCC TGC GGC TGG TCG CCG CGC GCG CCG AAA TGA TAC CCG CGA TCA CGA AAT ACA TGT
ala ala cys gly trp ser pro arg ala pro lys)OPA(tyr pro arg ser arg asn thr cys
181/61   211/71   ~SEQ ID NO: 467
CGG CGC TGG ACG TCG CCG TGC TGG CCA GCT CGA CCG GAC ACG ATG TGG AGG GGG CGC AGA
arg arg trp thr ser pro cys trp pro ala arg pro asp thr met trp arg gly arg arg
241/81   271/91
AAA ACT TCA CCG CCC GCA AGT ACG AGC TGC AGA CGC GAC TGG CCG ACA CCG ACG TCA TCG
lys thr ser pro pro ala ser thr ser cys arg arg asp trp pro thr pro thr ser ser
301/101   331/111
CAG ACG TGC GGT CGG GAG TGA ACA CGC TGC TCA ACG GCG GTC AGG CGC TGC TGG ATA AGA
gln thr cys gly arg glu)OPA(thr arg cys ser thr ala val arg arg cys trp ile arg
361/121   ~SEQ ID NO: 468
TGC TGG CCG ACA GCA TCG GCT TGC GGG ATC
cys trp pro thr ala ser ala cys gly ile)

SEQ ID NOS:463-468

FIG. 21A

___SEQ ID NO: 469   32/11
TCC TGG TCG CCG CGC AAC TGG CCG GTC CCG ATG GAA AGT GTT CAC GAT CGC GCT TCT GCC
(ser trp ser pro arg asn trp pro val pro met glu ser val his asp arg ala ser ala
62/21 ~SEQ ID NO: 470   92/31
GCT GGT AGT GGC GAT GGT GTT AGC AGG ATT GCG GGT CGA GGC TGC GAT GGC CAG CAC CAG
ala gly ser gly asp gly val ser arg ile ala gly arg gly cys asp gly gln his gln
122/41   152/51
CGG CCT GCG GCT GGT CGC CGC GCG CGC CGA AAT GAT ACC CGC GAT CAC GAA ATA CAT GTC
arg pro ala ala gly arg arg ala arg arg asn asp thr arg asp his glu ile his val
182/61   212/71
GGC GCT GGA CGT CGC CGT GCT GGC CAG CTC GAC CGG ACA CGA TGT GGA GGG GGC GCA GAA
gly ala gly arg arg arg ala gly gln leu asp arg thr arg cys gly gly gly ala glu
242/81   272/91
AAA CTT CAC CGC CCG CAA GTA CGA GCT GCA GAC GCG ACT GGC CGA CAC CGA CGT CAT CGC
lys leu his arg pro gln val arg ala ala asp ala thr gly arg his arg arg his arg
302/101   332/111
AGA CGT GCG GTC GGG AGT GAA CAC GCT GCT CAA CGG CGG TCA GGC GCT GCT GGA TAA GAT
arg arg ala val gly ser glu his ala ala gln arg arg ser gly ala ala gly)OCH(asp
362/121   SEQ ID NO: 471___
GCT GGC CGA CAG CAT CGG CTT GCG GGA TC
ala gly arg gln his arg leu ala gly)

SEQ ID NOS:469-471

FIG. 21B

SEQ ID NO: 472   33/11
CCT GGT CGC CGC GCA ACT GGC CGG TCC CGA TGG AAA GTG TTC ACG ATC GCG CTT CTG CCG
(pro gly arg arg ala thr gly arg ser arg trp lys val phe thr ile ala leu leu pro
63/21 SEEQ ID NO: 473   93/31
CTG GTA GTG GCG ATG GTG TTA GCA GGA TTG CGG GTC GAG GCT GCG ATG GCC AGC ACC AGC
leu val val ala met val leu ala gly leu arg val glu ala ala met ala ser thr ser
123/41   153/51
GGC CTG CGG CTG GTC GCC GCG CGC GCC GAA ATG ATA CCC GCG ATC ACG AAA TAC ATG TCG
gly leu arg leu val ala ala arg ala glu met ile pro ala ile thr lys tyr met ser
183/61   213/71
GCG CTG GAC GTC GCC GTG CTG GCC AGC TCG ACC GGA CAC GAT GTG GAG GGG GCG CAG AAA
ala leu asp val ala val leu ala ser ser thr gly his asp val glu gly ala gln lys
243/81   273/91
AAC TTC ACC GCC CGC AAG TAC GAG CTG CAG ACG CGA CTG GCC GAC ACC GAC GTC ATC GCA
asn phe thr ala arg lys tyr glu leu gln thr arg leu ala asp thr asp val ile ala
303/101   333/111
GAC GTG CGG TCG GGA GTG AAC ACG CTG CTC AAC GGC GGT CAG GCG CTG CTG GAT AAG ATG
asp val arg ser gly val asn thr leu leu asn gly gly gln ala leu leu asp lys met
363/121
CTG GCC GAC AGC ATC GGC TTG CGG GAT C
leu ala asp ser ile gly leu arg asp)

SEQ ID NOS:472-473

FIG. 21C

part of the nucleotide sequence of seq21A

1/1 SEQ ID NO: 474   31/11
ACG ATC GCG CTT CTG CCG CTG GTA GTG GCG ATG GTG TTA GCA GGA TTG CGG GTC GAG GCT
(thr ile ala leu leu pro leu val val ala met val leu ala gly leu arg val glu ala
61/21 SEEQ ID NO: 475   91/31
GCG ATG GCC AGC ACC AGC GGC CTG CGG CTG GTC GCC GCG CGC GCC GAA ATG ATA CCC GCG
ala met ala ser thr ser gly leu arg leu val ala ala arg ala glu met ile pro ala
121/41   151/51
ATC ACG AAA TAC ATG TCG GCG CTG GAC GTC GCC GTG CTG GCC AGC TCG ACC GGA CAC GAT
ile thr lys tyr met ser ala leu asp val ala val leu ala ser ser thr gly his asp
181/61   211/71
GTG GAG GGG GCG CAG AAA AAC TTC ACC GCC CGC AAG TAC GAG CTG CAG ACG CGA CTG GCC
val glu gly ala gln lys asn phe thr ala arg lys tyr glu leu gln thr arg leu ala
241/81   271/91
GAC ACC GAC GTC ATC GCA GAC GTG CGG TCG GGA GTG AAC ACG CTG CTC AAC GGC GGT CAG
asp thr asp val ile ala asp val arg ser gly val asn thr leu leu asn gly gly gln
301/101   331/111
GCG CTG CTG GAT AAG ATG CTG GCC GAC AGC ATC GGC TTG CGG GAT C
ala leu leu asp lys met leu ala asp ser ile gly leu arg asp)

SEQ ID NOS:474-475

FIG. 21A'

1/1 ——SEQ ID NO: 476   SEQ ID NO: 478 31/11
CGA TCG CGC TTC TGC CGC TGG TAG TGG CGA TGG TGT TAG CAG GAT TGC GGG TCG AGG CTG
(arg ser arg phe cys arg trp)AMB(trp arg trp cys)AMB(gln asp cys gly ser arg leu
61/21 ——SEQ ID NO: 477   91/31   ——SEQ ID NO: 479
CGA TGG CCA GCA CCA GCG GCC TGC GGC TGG TCG CCG CGC GCG CCG AAA TGA TAC CCG CGA
arg trp pro ala pro ala ala cys gly trp ser pro arg ala pro lys)OPA(tyr pro arg
121/41   151/51   SEQ ID NO: 480
TCA CGA AAT ACA TGT CGG CGC TGG ACG TCG CCG TGC TGG CCA GCT CGA CCG GAC ACG ATG
ser arg asn thr cys arg arg trp thr ser pro cys trp pro ala arg pro asp thr met
181/61   211/71
TGG AGG GGG CGC AGA AAA ACT TCA CCG CCC GCA AGT ACG AGC TGC AGA CGC GAC TGG CCG
trp arg gly arg arg lys thr ser pro pro ala ser thr ser cys arg arg asp trp pro
241/81   271/91
ACA CCG ACG TCA TCG CAG ACG TGC GGT CGG GAG TGA ACA CGC TGC TCA ACG GCG GTC AGG
thr pro thr ser ser gln thr cys gly arg glu)OPA(thr arg cys ser thr ala val arg
301/101   331/111   ——SEQ ID NO: 481
CGC TGC TGG ATA AGA TGC TGG CCG ACA GCA TCG GCT TGC GGG ATC
arg cys trp ile arg cys trp pro thr ala ser ala cys gly ile)

SEQ ID NOS:476-481

FIG. 21B'

1/1 ——SEQ ID NO: 482   31/11
CAC GAT CGC GCT TCT GCC GCT GGT AGT GGC GAT GGT GTT AGC AGG ATT GCG GGT CGA GGC
(his asp arg ala ser ala ala gly ser gly asp gly val ser arg ile ala gly arg gly
61/21 ——SEQ ID NO: 483   91/31
TGC GAT GGC CAG CAC CAG CGG CCT GCG GCT GGT CGC CGC GCG CGC CGA AAT GAT ACC CGC
cys asp gly gln his gln arg pro ala ala gly arg arg ala arg arg asn asp thr arg
121/41   151/51
GAT CAC GAA ATA CAT GTC GGC GCT GGA CGT CGC CGT GCT GGC CAG CTC GAC CGG ACA CGA
asp his glu ile his val gly ala gly arg arg arg ala gly gln leu asp arg thr arg
181/61   211/71
TGT GGA GGG GGC GCA GAA AAA CTT CAC CGC CCG CAA GTA CGA GCT GCA GAC GCG ACT GGC
cys gly gly gly ala glu lys leu his arg pro gln val arg ala ala asp ala thr gly
241/81   271/91
CGA CAC CGA CGT CAT CGC AGA CGT GCG GTC GGG AGT GAA CAC GCT GCT CAA CGG CGG TCA
arg his arg arg his arg arg arg ala val gly ser glu his ala ala gln arg arg ser
301/101   331/111
GGC GCT GCT GGA TAA GAT GCT GGC CGA CAG CAT CGG CTT GCG GGA TC
gly ala ala gly)OCH(asp ala gly arg gln his arg leu ala gly)
                      ——SEQ ID NO: 484

SEQ ID NOS:482-484

FIG. 21C'

sequence Rv3365c predicted by Cole et al. (Nature 393:537-544) and
containing Seq21A'

1/1 ──SEQ ID NO: 485   31/11
gtg acc atg ttc gcc cgc ccg acc atc ccg gtc gcg gcg gcc gct tct gat att tcc gcc
(val thr met phe ala arg pro thr ile pro val ala ala ala ala ser asp ile ser ala
61/21 ──SEQ ID NO: 486   91/31
ccg gct caa ccg gcc cgc ggc aaa cct cag caa cgc ccg ccg tcc tgg tcg ccg cgc aac
pro ala gln pro ala arg gly lys pro gln gln arg pro pro ser trp ser pro arg asn
121/41   151/51
tgg ccg gtc cga tgg aaa gtg ttc acg atc gcg ctt ctg ccg ctg gta gtg gcg atg gtg
trp pro val arg trp lys val phe thr ile ala leu leu pro leu val val ala met val
181/61   211/71
tta gca gga ttg cgg gtc gag gct gcg atg gcc agc acc agc ggc ctg cgg ctg gtc gcc
leu ala gly leu arg val glu ala ala met ala ser thr ser gly leu arg leu val ala
241/81   271/91
gcg cgc gcc gaa atg ata ccc gcg atc acg aaa tac atg tcg gcg ctg gac gtc gcc gtg
ala arg ala glu met ile pro ala ile thr lys tyr met ser ala leu asp val ala val
301/101   331/111
ctg gcc agc tcg acc gga cac gat gtg gag ggg gcg cag aaa aac ttc acc gcc cgc aag
leu ala ser ser thr gly his asp val glu gly ala gln lys asn phe thr ala arg lys
361/121   391/131
tac gag ctg cag acg cga ctg gcc gac acc gac gtc atc gca gac gtg cgg tcg gga gtg
tyr glu leu gln thr arg leu ala asp thr asp val ile ala asp val arg ser gly val
421/141   451/151
aac acg ctg ctc aac ggc ggt cag gcg ctg ctg gat aag gtg ctg gcc gac agc atc ggc
asn thr leu leu asn gly gly gln ala leu leu asp lys val leu ala asp ser ile gly
481/161   511/171
ttg cgg gat cgg gtc acc gcc tac gcg ccg ctg ctg ttg acg gcc cag aac gtg att gac
leu arg asp arg val thr ala tyr ala pro leu leu leu thr ala gln asn val ile asp
541/181   571/191
gcg tcg gtg cgg gtt gac agc gag caa atc cga acc cag gtg cag ggt ttg agc cga gcc
ala ser val arg val asp ser glu gln ile arg thr gln val gln gly leu ser arg ala
601/201   631/211
gtt ggc gcc cgc ggg cag atg acg atg cag gag atc ctg gtg act cgc ggc gcc gac ctt
val gly ala arg gly gln met thr met gln glu ile leu val thr arg gly ala asp leu
661/221   691/231
gcc gag ccg caa ctg cgc agc gcg atg gtt acc ctg gcc ggc acc gaa ccc tcg acg ctg
ala glu pro gln leu arg ser ala met val thr leu ala gly thr glu pro ser thr leu
721/241   751/251
ttc ggg atg agc gcg gcg ctc ggt gca ggc tcg ccg gac acc aag aac ctg cag cag caa
phe gly met ser ala ala leu gly ala gly ser pro asp thr lys asn leu gln gln gln
781/261   811/271
atg gtg acc agg atg gcg atc atg tcc gat ccg gcc gtt gca ctg gtc aac aac cca gag
met val thr arg met ala ile met ser asp pro ala val ala leu val asn asn pro glu
841/281   871/291
ctg ctg cac tcg ata cag atc acc cgc gac att gcc gag cag gtg atc acc gac acc acc
leu leu his ser ile gln ile thr arg asp ile ala glu gln val ile thr asp thr thr
901/301   931/311
gag gcg gtg acg aag tcg gtg caa agc cag gcc acc gac cgg cgg gat gcc gcg att cgc
glu ala val thr lys ser val gln ser gln ala thr asp arg arg asp ala ala ile arg
961/321   991/331
gac gcc gtg ctg gtg ttg gcc gcc atc gcg acc gcg atc gtc gtc gtg ttg gtg gtg gcg
asp ala val leu val leu ala ala ile ala thr ala ile val val val leu val val ala

SEQ ID NOS:485-486

FIG. 21D

1021/341
cgc acg ctg gtc ggg ccg atg cgg gta ctg
arg thr leu val gly pro met arg val leu
1081/361
gat ctc gac ggc gag atc gcg gcg gtc cgc
asp leu asp gly glu ile ala ala val arg
1141/381
gcg gtg tac acc acc gag gaa atc ggt cag
ala val tyr thr thr glu glu ile gly gln
1201/401
cgg gcc ctg ttg ctg gcc ggc gag gaa acg
arg ala leu leu leu ala gly glu glu thr
1261/421
gag acc atg tcg cgg cgt agc cgt tcc ctg
glu thr met ser arg arg ser arg ser leu
1321/441
ctg gag cgc aac gag gag gat ccc gcc cga
leu glu arg asn glu glu asp pro ala arg
1381/461
gcc gcc cgg ctg cgc cgc aac agc gcc aac
ala ala arg leu arg arg asn ser ala asn
1441/481
cgt gac cac cgc gag ccg gtg ccg ctg tca
arg asp his arg glu pro val pro leu ser
1501/501
gag gac tat cgc cgc gtc gac atc gcg agg
glu asp tyr arg arg val asp ile ala arg
1561/521
gct ggt ggc gtc att cat ctg ctt gcc gag
ala gly gly val ile his leu leu ala glu
1621/541
ccg acc aca ccc gtt cgg gtt gcc gcc gca
pro thr thr pro val arg val ala ala ala
1681/561
atc tcg gat tcc ggc ctg ggc atg acc gat
ile ser asp ser gly leu gly met thr asp
1741/581
cgg gcc ggc ggt gag gtc acc ccg gat agt
arg ala gly gly glu val thr pro asp ser
1801/601
cgg ctg gcc ggt cgg cac ggc atc cga gtc
arg leu ala gly arg his gly ile arg val
1861/621
ggc acc ggc acc acc gcc gag gtc tac ctg
gly thr gly thr thr ala glu val tyr leu
1921/641
gcg cag ccg cca aag ccg cgg gta ttt gcg
ala gln pro pro lys pro arg val phe ala
1981/661
gcc gat ccg acg gac gtt ccc gcc gcc atc
ala asp pro thr asp val pro ala ala ile

1051/351
cgt gat ggg gcg ctc aag gtt gct cat acc
arg asp gly ala leu lys val ala his thr
1111/371
gcc ggc gac gag ccg atc ccc gag cca ctg
ala gly asp glu pro ile pro glu pro leu
1171/391
gtc gcg cat gcg gtc gac gag ctg cac acc
val ala his ala val asp glu leu his thr
1231/411
cgg ttg cga ctg ctg gtc aac gag atg ttt
arg leu arg leu leu val asn glu met phe
1291/431
gtc gac cag cag ctg tcg gtc atc gac caa
val asp gln gln leu ser val ile asp gln
1351/451
ctc gac agc ctt ttc cgg ctc gat cac ctg
leu asp ser leu phe arg leu asp his leu
1411/471
ctg ctg gtg ctg gcc ggt gcg cag att acc
leu leu val leu ala gly ala gln ile thr
1471/491
acc gtg atc agc gcc gcc gtg tca gag gtc
thr val ile ser ala ala val ser glu val
1531/511
gta ccc gac tgt gcg gta gtc ggc gca gcg
val pro asp cys ala val val gly ala ala
1591/531
ctg atc gac aac gcg ttg cgc tac tcg tca
leu ile asp asn ala leu arg tyr ser ser
1651/551
atc ggc agc gaa ggc agt gtt ctg ctg cga
ile gly ser glu gly ser val leu leu arg
1711/571
gcc gat cgg cgg atg gcc aat atg cgg ctg
ala asp arg arg met ala asn met arg leu
1771/591
gcc cgg cac atg ggt ctg ttc gta gtc ggc
ala arg his met gly leu phe val val gly
1831/611
ggg ctg cgc ggt ccg gtg acc ggt gaa cag
gly leu arg gly pro val thr gly glu gln
1891/631
ccg cta gcc gtg ctc gag ggg acg gcc cca
pro leu ala val leu glu gly thr ala pro
1951/651
atc aag ccg ccg tgt cct gaa ccc gcg gcg
ile lys pro pro cys pro glu pro ala ala
2011/671
ggg ccg cta cca ccg gtc acg ttg ctc ccg
gly pro leu pro pro val thr leu leu pro

SEQ ID NOS:485-486 (continued 1)

FIG. 21D (continued 1)

```
2041/681                                      2071/691
cgc cgt acc ccg ggg tcc agt ggc atc gcc      gac gtc ccg gcc cag ccg atg cag cag cgg
arg arg thr pro gly ser ser gly ile ala      asp val pro ala gln pro met gln gln arg
2101/701                                      2131/711
cgg cgc gag ctg aaa aca ccc tgg tgg gag      gat agg ttt caa cag gag ccc aaa caa ccg
arg arg glu leu lys thr pro trp trp glu      asp arg phe gln gln glu pro lys gln pro
2161/721                                      2191/731
ccc gca cca gaa ccg cga ccg gcg ccg ccg      ccc gcc aaa ccc gcg cca ccg gcg ggc ccg
pro ala pro glu pro arg pro ala pro pro      pro ala lys pro ala pro pro ala gly pro
2221/741                                      2251/751
gtt gat gac gac gtc atc tac cgg cgg atg      ctc tcc gag atg gtg ggt gac ccg cac gag
val asp asp asp val ile tyr arg arg met      leu ser glu met val gly asp pro his glu
2281/761                                      2311/771
ctg gcc cac agc ccc gat ctg gac tgg aag      tcg gtg tgg gac cac ggc tgg tcg gcg gcc
leu ala his ser pro asp leu asp trp lys      ser val trp asp his gly trp ser ala ala
2341/781                                      2371/791
gcc gag gcc gcg gac aag ccc gtg cag tcc      cgc acg gac tac ggc ctg ccg gtg cgc gaa
ala glu ala ala asp lys pro val gln ser      arg thr asp tyr gly leu pro val arg glu
2401/801                                      2431/811
ccc ggg gcc cgg tta gtg ccg ggg gcg gcg      gtg cct gag gga ccc gat cgg gag cat ccg
pro gly ala arg leu val pro gly ala ala      val pro glu gly pro asp arg glu his pro
2461/821                                      2491/831
ggt gca gcg cta gca tcc aac ggc gga ctt      cat ccc ggc cga gcg ccg cgg cac gcg gct
gly ala ala leu ala ser asn gly gly leu      his pro gly arg ala pro arg his ala ala
2521/841                                      2551/851
gcg gta cgc gac ccc gac gcg gtt cgt gcc      tcc atc agc agc cat ttc ggc ggc gtg cgc
ala val arg asp pro asp ala val arg ala      ser ile ser ser his phe gly gly val arg
2581/861                                      2611/871
acc ggg cgg tcg cat gcc cgc gag agc agt      cag gga ccc aat cag caa tga
thr gly arg ser his ala arg glu ser ser      gln gly pro asn gln gln)OPA
```

SEQ ID NOS:485-486 (continued)

FIG. 21D (continued)

ORF according to Cole et al. (Nature 393:537-544) and containing Rv3365c

```
1/1    SEQ ID NO: 487                    31/11
taa ggg tgc ggc cgg tgg cac ggc cgc ggc cac gtg acc atg ttc gcc cgc ccg acc atc
OCH(gly cys gly arg trp his gly arg gly his val thr met phe ala arg pro thr ile
61/21    SEQ ID NO: 488                  91/31
ccg gtc gcg gcg gcc gct tct gat att tcc gcc ccg gct caa ccg gcc cgc ggc aaa cct
pro val ala ala ala ala ser asp ile ser ala pro ala gln pro ala arg gly lys pro
121/41                                   151/51
cag caa cgc ccg ccg tcc tgg tcg ccg cgc aac tgg ccg gtc cga tgg aaa gtg ttc acg
gln gln arg pro pro ser trp ser pro arg asn trp pro val arg trp lys val phe thr
181/61                                   211/71
atc gcg ctt ctg ccg ctg gta gtg gcg atg gtg tta gca gga ttg cgg gtc gag gct gcg
ile ala leu leu pro leu val val ala met val leu ala gly leu arg val glu ala ala
241/81                                   271/91
atg gcc agc acc agc ggc ctg cgg ctg gtc gcc gcg cgc gcc gaa atg ata ccc gcg atc
met ala ser thr ser gly leu arg leu val ala ala arg ala glu met ile pro ala ile
301/101                                  331/111
acg aaa tac atg tcg gcg ctg gac gtc gcc gtg ctg gcc agc tcg acc gga cac gat gtg
thr lys tyr met ser ala leu asp val ala val leu ala ser ser thr gly his asp val
361/121                                  391/131
gag ggg gcg cag aaa aac ttc acc gcc cgc aag tac gag ctg cag acg cga ctg gcc gac
glu gly ala gln lys asn phe thr ala arg lys tyr glu leu gln thr arg leu ala asp
421/141                                  451/151
acc gac gtc atc gca gac gtg cgg tcg gga gtg aac acg ctg ctc aac ggc ggt cag gcg
thr asp val ile ala asp val arg ser gly val asn thr leu leu asn gly gly gln ala
481/161                                  511/171
ctg ctg gat aag gtg ctg gcc gac agc atc ggc ttg cgg gat cgg gtc acc gcc tac gcg
leu leu asp lys val leu ala asp ser ile gly leu arg asp arg val thr ala tyr ala
541/181                                  571/191
ccg ctg ctg ttg acg gcc cag aac gtg att gac gcg tcg gtg cgg gtt gac agc gag caa
pro leu leu leu thr ala gln asn val ile asp ala ser val arg val asp ser glu gln
601/201                                  631/211
atc cga acc cag gtg cag ggt ttg agc cga gcc gtt ggc gcc cgc ggg cag atg acg atg
ile arg thr gln val gln gly leu ser arg ala val gly ala arg gly gln met thr met
661/221                                  691/231
cag gag atc ctg gtg act cgc ggc gcc gac ctt gcc gag ccg caa ctg cgc agc gcg atg
gln glu ile leu val thr arg gly ala asp leu ala glu pro gln leu arg ser ala met
721/241                                  751/251
gtt acc ctg gcc ggc acc gaa ccc tcg acg ctg ttc ggg atg agc gcg gcg ctc ggt gca
val thr leu ala gly thr glu pro ser thr leu phe gly met ser ala ala leu gly ala
781/261                                  811/271
ggc tcg ccg gac acc aag aac ctg cag cag caa atg gtg acc agg atg gcg atc atg tcc
gly ser pro asp thr lys asn leu gln gln gln met val thr arg met ala ile met ser
841/281                                  871/291
gat ccg gcc gtt gca ctg gtc aac aac cca gag ctg ctg cac tcg ata cag atc acc cgc
asp pro ala val ala leu val asn asn pro glu leu leu his ser ile gln ile thr arg
```

SEQ ID NOS:487-488

FIG. 21E

901/301
gac att gcc gag cag gtg atc acc gac acc
asp ile ala glu gln val ile thr asp thr
961/321
cag gcc acc gac cgg cgg gat gcc gcg att
gln ala thr asp arg arg asp ala ala ile
1021/341
gcg acc gcg atc gtc gtc gtg ttg gtg gtg
ala thr ala ile val val val leu val val
1081/361
ctg cgt gat ggg gcg ctc aag gtt gct cat
leu arg asp gly ala leu lys val ala his
1141/381
cgc gcc ggc gac gag ccg atc ccc gag cca
arg ala gly asp glu pro ile pro glu pro
1201/401
cag gtc gcg cat gcg gtc gac gag ctg cac
gln val ala his ala val asp glu leu his
1261/421
acg cgg ttg cga ctg ctg gtc aac gag atg
thr arg leu arg leu leu val asn glu met
1321/441
ctg gtc gac cag cag ctg tcg gtc atc gac
leu val asp gln gln leu ser val ile asp
1381/461
cga ctc gac agc ctt ttc cgg ctc gat cac
arg leu asp ser leu phe arg leu asp his
1441/481
aac ctg ctg gtg ctg gcc ggt gcg cag att
asn leu leu val leu ala gly ala gln ile
1501/501
tca acc gtg atc agc gcc gcc gtg tca gag
ser thr val ile ser ala ala val ser glu
1561/521
agg gta ccc gac tgt gcg gta gtc ggc gca
arg val pro asp cys ala val val gly ala
1621/541
gag ctg atc gac aac gcg ttg cgc tac tcg
glu leu ile asp asn ala leu arg tyr ser
1681/561
gca atc ggc agc gaa ggc agt gtt ctg ctg
ala ile gly ser glu gly ser val leu leu
1741/581
gat gcc gat cgg cgg atg gcc aat atg cgg
asp ala asp arg arg met ala asn met arg
1801/601
agt gcc cgg cac atg ggt ctg ttc gta gtc
ser ala arg his met gly leu phe val val

931/311
acc gag gcg gtg acg aag tcg gtg caa agc
thr glu ala val thr lys ser val gln ser
991/331
cgc gac gcc gtg ctg gtg ttg gcc gcc atc
arg asp ala val leu val leu ala ala ile
1051/351
gcg cgc acg ctg gtc ggg ccg atg cgg gta
ala arg thr leu val gly pro met arg val
1111/371
acc gat ctc gac ggc gag atc gcg gcg gtc
thr asp leu asp gly glu ile ala ala val
1171/391
ctg gcg gtg tac acc acc gag gaa atc ggt
leu ala val tyr thr thr glu glu ile gly
1231/411
acc cgg gcc ctg ttg ctg gcc ggc gag gaa
thr arg ala leu leu leu ala gly glu glu
1291/431
ttt gag acc atg tcg cgg cgt agc cgt tcc
phe glu thr met ser arg arg ser arg ser
1351/451
caa ctg gag cgc aac gag gag gat ccc gcc
gln leu glu arg asn glu glu asp pro ala
1411/471
ctg gcc gcc cgg ctg cgc cgc aac agc gcc
leu ala ala arg leu arg arg asn ser ala
1471/491
acc cgt gac cac cgc gag ccg gtg ccg ctg
thr arg asp his arg glu pro val pro leu
1531/511
gtc gag gac tat cgc cgc gtc gac atc gcg
val glu asp tyr arg arg val asp ile ala
1591/531
gcg gct ggt ggc gtc att cat ctg ctt gcc
ala ala gly gly val ile his leu leu ala
1651/551
tca ccg acc aca ccc gtt cgg gtt gcc gcc
ser pro thr thr pro val arg val ala ala
1711/571
cga atc tcg gat tcc ggc ctg ggc atg acc
arg ile ser asp ser gly leu gly met thr
1771/591
ctg cgg gcc ggc ggt gag gtc acc ccg gat
leu arg ala gly gly glu val thr pro asp
1831/611
ggc cgg ctg gcc ggt cgg cac ggc atc cga
gly arg leu ala gly arg his gly ile arg

SEQ ID NOS:487-488 (continued 1)

FIG. 21E (continued 1)

```
1861/621                                      1891/631
gtc ggg ctg cgc ggt ccg gtg acc ggt gaa      cag ggc acc ggc acc acc gcc gag gtc tac
val gly leu arg gly pro val thr gly glu      gln gly thr gly thr thr ala glu val tyr
1921/641                                      1951/651
ctg ccg cta gcc gtg ctc gag ggg acg gcc      cca gcg cag ccg cca aag ccg cgg gta ttt
leu pro leu ala val leu glu gly thr ala      pro ala gln pro pro lys pro arg val phe
1981/661                                      2011/671
gcg atc aag ccg ccg tgt cct gaa ccc gcg      gcg gcc gat ccg acg gac gtt ccc gcc gcc
ala ile lys pro pro cys pro glu pro ala      ala ala asp pro thr asp val pro ala ala
2041/681                                      2071/691
atc ggg ccg cta cca ccg gtc acg ttg ctc      ccg cgc cgt acc ccg ggg tcc agt ggc atc
ile gly pro leu pro pro val thr leu leu      pro arg arg thr pro gly ser ser gly ile
2101/701                                      2131/711
gcc gac gtc ccg gcc cag ccg atg cag cag      cgg cgg cgc gag ctg aaa aca ccc tgg tgg
ala asp val pro ala gln pro met gln gln      arg arg arg glu leu lys thr pro trp trp
2161/721                                      2191/731
gag gat agg ttt caa cag gag ccc aaa caa      ccg ccc gca cca gaa ccg cga ccg gcg ccg
glu asp arg phe gln gln glu pro lys gln      pro pro ala pro glu pro arg pro ala pro
2221/741                                      2251/751
ccg ccc gcc aaa ccc gcg cca ccg gcg ggc      ccg gtt gat gac gac gtc atc tac cgg cgg
pro pro ala lys pro ala pro pro ala gly      pro val asp asp asp val ile tyr arg arg
2281/761                                      2311/771
atg ctc tcc gag atg gtg ggt gac ccg cac      gag ctg gcc cac agc ccc gat ctg gac tgg
met leu ser glu met val gly asp pro his      glu leu ala his ser pro asp leu asp trp
2341/781                                      2371/791
aag tcg gtg tgg gac cac ggc tgg tcg gcg      gcc gcc gag gcc gcg gac aag ccc gtg cag
lys ser val trp asp his gly trp ser ala      ala ala glu ala ala asp lys pro val gln
2401/801                                      2431/811
tcc cgc acg gac tac ggc ctg ccg gtg cgc      gaa ccc ggg gcc cgg tta gtg ccg ggg gcg
ser arg thr asp tyr gly leu pro val arg      glu pro gly ala arg leu val pro gly ala
2461/821                                      2491/831
gcg gtg cct gag gga ccc gat cgg gag cat      ccg ggt gca gcg cta gca tcc aac ggc gga
ala val pro glu gly pro asp arg glu his      pro gly ala ala leu ala ser asn gly gly
2521/841                                      2551/851
ctt cat ccc ggc cga gcg ccg cgg cac gcg      gct gcg gta cgc gac ccc gac gcg gtt cgt
leu his pro gly arg ala pro arg his ala      ala ala val arg asp pro asp ala val arg
2581/861                                      2611/871
gcc tcc atc agc agc cat ttc ggc ggc gtg      cgc acc ggg cgg tcg cat gcc cgc gag agc
ala ser ile ser ser his phe gly gly val      arg thr gly arg ser his ala arg glu ser
2641/881
agt cag gga ccc aat cag caa tga
ser gln gly pro asn gln gln)OPA
```

SEQ ID NOS:487-488 (continued 2)

FIG. 21E (continued 2)

SEQ ID NO: 489   31/11
CTA CGA CAA GGC AAA GGA GCA CAG GGT GAA GCG TGG ACT GAC GGT CGC GGT AGC CGG AGC
(leu arg gln gly lys gly ala gln gly glu ala trp thr asp gly arg gly ser arg ser
61/21   SEQ ID NO: 490   91/31
CGC CAT TCT GGT CGC AGG TCT TTC CGG ATG TTC AAG CAA CAA GTC GAC TAC AGG AAG CGG
arg his ser gly arg arg ser phe arg met phe lys gln gln val asp tyr arg lys arg)
121/41   151/51
TGA GAC CAC GAC CGC GGC AGG CAC GAC GGC AAG CCC CGG CGC CGC ATC CGG GCC GAA GGT
OPA(asp his asp arg gly arg his asp gly lys pro arg arg arg ile arg ala glu gly
181/61   SEQ ID NO: 491   211/71
CGT CAT CGA CGG TAA GGA CCA GAA CGT CAC CGG GTC TGT GGT GTG CAC AAC CGC GGC CGG
arg his arg arg)OCH(gly pro glu arg his arg val cys gly val his asn arg gly arg
241/81   SEQ ID NO: 492   271/91
CAA TGT CAA CAT CGC GAT CGG CGG GGC GGC GAC CGG CAT TGC CGC CGT GCT CAC CGA CGG
gln cys gln his arg asp arg arg gly gly asp arg his cys arg arg ala his arg arg
301/101   331/111
CAA CCC TCC GGA GGT GAA GTC CGT TGG GCT CGG TAA CGT CAA CGG CGT CAC GCT GGG ATA
gln pro ser gly gly glu val arg trp ala arg)OCH(arg gln arg arg his ala gly ile
361/121   391/131 SEQ ID NO: 493
CAC GTC GGG CAC CGG ACA GGG TAA CGC TCG GCA ACC AAG GAC GGC AGC CAC TAC AAG ATC
his val gly his arg thr gly)OCH(arg ser ala thr lys asp gly ser his tyr lys ile)
                                   SEQ ID NO: 494

                           SEQ ID NOS:489-494

FIG. 22A

SEQ ID NO: 495   32/11
TAC GAC AAG GCA AAG GAG CAC AGG GTG AAG CGT GGA CTG ACG GTC GCG GTA GCC GGA GCC
(tyr asp lys ala lys glu his arg val lys arg gly leu thr val ala val ala gly ala
62/21   SEQ ID NO: 496   92/31
GCC ATT CTG GTC GCA GGT CTT TCC GGA TGT TCA AGC AAC AAG TCG ACT ACA GGA AGC GGT
ala ile leu val ala gly leu ser gly cys ser ser asn lys ser thr thr gly ser gly
122/41   152/51
GAG ACC ACG ACC GCG GCA GGC ACG ACG GCA AGC CCC GGC GCC GCA TCC GGG CCG AAG GTC
glu thr thr thr ala ala gly thr thr ala ser pro gly ala ala ser gly pro lys val
182/61   212/71
GTC ATC GAC GGT AAG GAC CAG AAC GTC ACC GGG TCT GTG GTG TGC ACA ACC GCG GCC GGC
val ile asp gly lys asp gln asn val thr gly ser val val cys thr thr ala ala gly
242/81   272/91
AAT GTC AAC ATC GCG ATC GGC GGG GCG GCG ACC GGC ATT GCC GCC GTG CTC ACC GAC GGC
asn val asn ile ala ile gly gly ala ala thr gly ile ala ala val leu thr asp gly
302/101   332/111
AAC CCT CCG GAG GTG AAG TCC GTT GGG CTC GGT AAC GTC AAC GGC GTC ACG CTG GGA TAC
asn pro pro glu val lys ser val gly leu gly asn val asn gly val thr leu gly tyr
362/121   392/131
ACG TCG GGC ACC GGA CAG GGT AAC GCT CGG CAA CCA AGG ACG GCA GCC ACT ACA AGA TC
thr ser gly thr gly gln gly asn ala arg gln pro arg thr ala ala thr thr arg)

                           SEQ ID NOS:495-496

FIG. 22B

SEQ ID NO: 497   33/11
ACG ACA AGG CAA AGG AGC ACA GGG TGA AGC GTG GAC TGA CGG TCG CGG TAG CCG GAG CCG
(thr thr arg gln arg ser thr gly)OPA ser val asp OPA arg ser arg AMB(pro glu pro
63/21   SEQ ID NO: 498   93/31   SEQ ID NO: 499
CCA TTC TGG TCG CAG GTC TTT CCG GAT GTT CAA GCA ACA AGT CGA CTA CAG GAA GCG GTG
pro phe trp ser gln val phe pro asp val gln ala thr ser arg leu gln glu ala val
123/41   153/51
AGA CCA CGA CCG CGG CAG GCA CGA CGG CAA GCC CCG GCG CCG CAT CCG GGC CGA AGG TCG
arg pro arg pro arg gln ala arg arg gln ala pro ala pro his pro gly arg arg ser
183/61   213/71
TCA TCG ACG GTA AGG ACC AGA ACG TCA CCG GGT CTG TGG TGT GCA CAA CCG CGG CCG GCA
ser ser thr val arg thr arg thr ser pro gly leu trp cys ala gln pro arg pro ala
243/81   273/91
ATG TCA ACA TCG CGA TCG GCG GGG CGG CGA CCG GCA TTG CCG CCG TGC TCA CCG ACG GCA
met ser thr ser arg ser ala gly arg arg pro ala leu pro pro cys ser pro thr ala
303/101   333/111
ACC CTC CGG AGG TGA AGT CCG TTG GGC TCG GTA ACG TCA ACG GCG TCA CGC TGG GAT ACA
thr leu arg arg)OPA(ser pro leu gly ser val thr ser thr ala ser arg trp asp thr
363/121   SEQ ID NO: 500   393/131
CGT CGG GCA CCG GAC AGG GTA ACG CTC GGC AAC CAA GGA CGG CAG CCA CTA CAA GAT C
arg arg ala pro asp arg val thr leu gly asn gln gly arg gln pro leu gln asp)

SEQ ID NOS:497-500

FIG. 22C

SEQ ID NO: 501   31/11
GCA CAA CCG CGG CCG GCA ATG TCA ACA TCG CGA TCG GCG GGG CGG CGA CCG GCA TTG CCG
(ala gln pro arg pro ala met ser thr ser arg ser ala gly arg arg pro ala leu pro
61/21   SEQ ID NO: 502   91/31
CCG TGC TCA CCG ACG GCA ACC CTC CGG AGG TGA AGT CCG TTG GGC TCG GTA ACG TCA ACG
pro cys ser pro thr ala thr leu arg arg)OPA(ser pro leu gly ser val thr ser thr
121/41   151/51   SEQ ID NO: 503
GCG TCA CGC TGG GAT ACA CGT CGG GCA CCG GAC AGG GTA ACG CCT CGG CAA CCA AGG ACG
ala ser arg trp asp thr arg arg ala pro asp arg val thr pro arg gln pro arg thr
181/61   211/71
GCA GCC ACT ACA AGA TCA CAG GGT GAA GCG TGG ACT GAC GGT CGC GGT AGC CGG AGC CGC
ala ala thr thr arg ser gln gly glu ala trp thr asp gly arg gly ser arg ser arg
241/81   271/91
CAT TCT GGT CGC AGG TCT TTC CGG ATG TTC AAG CAA CAA GTC GAC TAC AGG AAG CGG TGA
his ser gly arg arg ser phe arg met phe lys gln gln val asp tyr arg lys arg)OPA
301/101   331/111
GAC CAC GAC CGC GGC AGG CAC GAC GGC AAG CCC CGG CGC CGC TCC GGG CCG AAG GTC GTC
(asp his asp arg gly arg his asp gly lys pro arg arg arg ser gly pro lys val val
361/121 SEQ ID NO: 504   391/131
ATC GAC GGT AAG GAC CAG AAC GTC ACC GGC TCC GTG GTG TGC ACA ACC GCG GCC GGC AAT
ile asp gly lys asp gln asn val thr gly ser val val cys thr thr ala ala gly asn
421/141   451/151
GTC AAC ATC GCG ATC GGC GGG GCG GCG ACC GGC ATT GCC GCC GTG CTC ACC GAC GGC AAC
val asn ile ala ile gly gly ala ala thr gly ile ala ala val leu thr asp gly asn
481/161   511/171
CCT CCG GAG GTG AAG TCC GTT GGG CTC GGT AAC GTC AAC GGC GTC ACG CTG GGA TAC ACG
pro pro glu val lys ser val gly leu gly asn val asn gly val thr leu gly tyr thr
541/181   571/191
TCG GGC ACC GGA CAG GGT AAC GCC TCG GCA ACC AAG GAC GGC AGC CAC TAC AAG ATC
ser gly thr gly gln gly asn ala ser ala thr lys asp gly ser his tyr lys ile)

SEQ ID NOS:501-504

FIG. 23A

SEQ ID NO: 505   32/11
CAC AAC CGC GGC CGG CAA TGT CAA CAT CGC GAT CGG CGG GGC GGC GAC CGG CAT TGC CGC
(his asn arg gly arg gln cys gln his arg asp arg arg gly gly asp arg his cys arg
62/21 SEQ ID NO: 506   92/31
CGT GCT CAC CGA CGG CAA CCC TCC GGA GGT GAA GTC CGT TGG GCT CGG TAA CGT CAA CGG
arg ala his arg arg gln pro ser gly gly glu val arg trp ala arg)OCH(arg gln arg
122/41   152/51 SEQ ID NO: 507
CGT CAC GCT GGG ATA CAC GTC GGG CAC CGG ACA GGG TAA CGC CTC GGC AAC CAA GGA CGG
arg his ala gly ile his val gly his arg thr gly)OCH(arg leu gly asn gln gly arg
182/61   212/71 SEQ ID NO: 508
CAG CCA CTA CAA GAT CAC AGG GTG AAG CGT GGA CTG ACG GTC GCG GTA GCC GGA GCC GCC
gln pro leu gln asp his arg val lys arg gly leu thr val ala val ala gly ala ala
242/81   272/91
ATT CTG GTC GCA GGT CTT TCC GGA TGT TCA AGC AAC AAG TCG ACT ACA GGA AGC GGT GAG
ile leu val ala gly leu ser gly cys ser ser asn lys ser thr thr gly ser gly glu
302/101   332/111
ACC ACG ACC GCG GCA GGC ACG ACG GCA AGC CCC GGC GCC GCT CCG GGC CGA AGG TCG TCA
thr thr thr ala ala gly thr thr ala ser pro gly ala ala pro gly arg arg ser ser
362/121   392/131
TCG ACG GTA AGG ACC AGA ACG TCA CCG GCT CCG TGG TGT GCA CAA CCG CGG CCG GCA ATG
ser thr val arg thr arg thr ser pro ala pro trp cys ala gln pro arg pro ala met
422/141   452/151
TCA ACA TCG CGA TCG GCG GGG CGG CGA CCG GCA TTG CCG CCG TGC TCA CCG ACG GCA ACC
ser thr ser arg ser ala gly arg arg pro ala leu pro pro cys ser pro thr ala thr
482/161   512/171
CTC CGG AGG TGA AGT CCG TTG GGC TCG GTA ACG TCA ACG GCG TCA CGC TGG GAT ACA CGT
leu arg arg)OPA(ser pro leu gly ser val thr ser thr ala ser arg trp asp thr arg
542/181   SEQ ID NO: 509 572/191
CGG GCA CCG GAC AGG GTA ACG CCT CGG CAA CCA AGG ACG GCA GCC ACT ACA AGA TC
arg ala pro asp arg val thr pro arg gln pro arg thr ala ala thr thr arg)

SEQ ID NOS:505-509

FIG. 23B

SEQ ID NO: 510   33/11
ACA ACC GCG GCC GGC AAT GTC AAC ATC GCG ATC GGC GGG GCG GCG ACC GGC ATT GCC GCC
(thr thr ala ala gly asn val asn ile ala ile gly gly ala ala thr gly ile ala ala
63/21   SEQ ID NO: 511   93/31
GTG CTC ACC GAC GGC AAC CCT CCG GAG GTG AAG TCC GTT GGG CTC GGT AAC GTC AAC GGC
val leu thr asp gly asn pro pro glu val lys ser val gly leu gly asn val asn gly
123/41   153/51
GTC ACG CTG GGA TAC ACG TCG GGC ACC GGA CAG GGT AAC GCC TCG GCA ACC AAG GAC GGC
val thr leu gly tyr thr ser gly thr gly gln gly asn ala ser ala thr lys asp gly
183/61   213/71
AGC CAC TAC AAG ATC ACA GGG TGA AGC GTG GAC TGA CGG TCG CGG TAG CCG CAG CCG CCA
ser his tyr lys ile thr gly)OPA ser val asp OPA arg ser arg AMB(pro glu pro pro
243/81   273/91   SEQ ID NO: 512
TTC TGG TCG CAG GTC TTT CCG GAT GTT CAA GCA ACA AGT CGA CTA CAG GAA GCG GTG AGA
phe trp ser gln val phe pro asp val gln ala thr ser arg leu gln glu ala val arg
303/101   333/111
CCA CGA CCG CGG CAG GCA CGA CGG CAA GCC CCG GCG CCG CTC CGG GCC GAA GGT CGT CAT
pro arg pro arg gln ala arg arg gln ala pro ala pro leu arg ala glu gly arg his
363/121   393/131
CGA CGG TAA GGA CCA GAA CGT CAC CGG CTC CGT GGT GTG CAC AAC CGC GGC CGG CAA TGT
arg arg)OCH(gly pro glu arg his arg leu arg gly val his asn arg gly arg gln cys
423/141   SEQ ID NO: 513   453/151
CAA CAT CGC GAT CGG CGG GGC GGC GAC CGG CAT TGC CGC CGT GCT CAC CGA CGG CAA CCC
gln his arg asp arg arg gly gly asp arg his cys arg arg ala his arg arg gln pro
483/161   513/171
TCC GGA GGT GAA GTC CGT TGG GCT CGG TAA CGT CAA CGG CGT CAC GCT GGG ATA CAC GTC
ser gly gly glu val arg trp ala arg)OCH(arg gln arg arg his ala gly ile his val
543/181   SEQ ID NO: 514   573/191
GGG CAC CGG ACA GGG TAA CGC CTC GGC AAC CAA GGA CGG CAG CCA CTA CAA GAT C
gly his arg thr gly)OCH(arg leu gly asn gln gly arg gln pro leu gln asp)
                 SEQ ID NO: 515

SEQ ID NOS:510-515

FIG. 23C

SEQ ID NO: 516   31/11
CTA ACG ACA GGC AAA GGA GCA CAG GGT GAA   GCG TGG ACT GAC GGT CGC GGT AGC CGG AGC
(leu thr thr gly lys gly ala gln gly glu   ala trp thr asp gly arg gly ser arg ser
61/21   SEQ ID NO: 517   91/31
CGC CAT TCT GGT CGC AGG TCT TTC CGG ATG   TTC AAG CAA CAA GTC GAC TAC AGG AAG CGG
arg his ser gly arg arg ser phe arg met   phe lys gln gln val asp tyr arg lys arg)
121/41   151/51
TGA GAC CAC GAC CGC GGC AGG CAC GAC GGC   AAG CCC CGG CGC CGC TCC GGG CCG AAG GTC
OPA(asp his asp arg gly arg his asp gly   lys pro arg arg arg ser gly pro lys val
181/61   SEQ ID NO: 518   211/71
GTC ATC GAC GGT AAG GAC CAG AAC GTC ACC   GGC TCC GTG GTG TGC ACA ACC GCG GCC GGC
val ile asp gly lys asp gln asn val thr   gly ser val val cys thr thr ala ala gly
241/81   271/91
AAT GTC AAC ATC GCG ATC GGC GGG GCG GCG   ACC GGC ATT GCC GCC GTG CTC ACC GAC GGC
asn val asn ile ala ile gly gly ala ala   thr gly ile ala ala val leu thr asp gly
301/101   331/111
AAC CCT CCG GAG GTG AAG TCC GTT GGG CTC   GGT AAC GTC AAC GGC GTC ACG CTG GGA TAC
asn pro pro glu val lys ser val gly leu   gly asn val asn gly val thr leu gly tyr
361/121   391/131
ACG TCG GGC ACC GGA CAG GGT AAC GCC TCG   GCA ACC AAG GAC GGC AGC CAC TAC AAG ATC
thr ser gly thr gly gln gly asn ala ser   ala thr lys asp gly ser his tyr lys ile)

SEQ ID NOS:516-518

FIG. 24A

SEQ ID NO: 519   32/11
TAA CGA CAG GCA AAG GAG CAC AGG GTG AAG   CGT GGA CTG ACG GTC GCG GTA GCC GGA GCC
OCH(arg gln ala lys glu his arg val lys   arg gly leu thr val ala val ala gly ala
62/21   SEQ ID NO: 520   92/31
GCC ATT CTG GTC GCA GGT CTT TCC GGA TGT   TCA AGC AAC AAG TCG ACT ACA GGA AGC GGT
ala ile leu val ala gly leu ser gly cys   ser ser asn lys ser thr thr gly ser gly
122/41   152/51
GAG ACC ACG ACC GCG GCA GGC ACG ACG GCA   AGC CCC GGC GCC GCT CCG GGC CGA AGG TCG
glu thr thr thr ala ala gly thr thr ala   ser pro gly ala ala pro gly arg arg ser
182/61   212/71
TCA TCG ACG GTA AGG ACC AGA ACG TCA CCG   GCT CCG TGG TGT GCA CAA CCG CGG CCG GCA
ser ser thr val arg thr arg thr ser pro   ala pro trp cys ala gln pro arg pro ala
242/81   272/91
ATG TCA ACA TCG CGA TCG GCG GGG CGG CGA   CCG GCA TTG CCG CCG TGC TCA CCG ACG GCA
met ser thr ser arg ser ala gly arg arg   pro ala leu pro pro cys ser pro thr ala
302/101   332/111
ACC CTC CGG AGG TGA AGT CCG TTG GGC TCG   GTA ACG TCA ACG GCG TCA CGC TGG GAT ACA
thr leu arg arg)OPA(ser pro leu gly ser   val thr ser thr ala ser arg trp asp thr
362/121   SEQ ID NO: 521   392/131
CGT CGG GCA CCG GAC AGG GTA ACG CCT CGG   CAA CCA AGG ACG GCA GCC ACT ACA AGA TC
arg arg ala pro asp arg val thr pro arg   gln pro arg thr ala ala thr thr arg)

SEQ ID NOS:519-521

FIG. 24B

SEQ ID NO: 522   33/11
AAC GAC AGG CAA AGG AGC ACA GGG TGA AGC GTG GAC TGA CGG TCG CGG TAG CCG GAG CCG
(asn asp arg gln arg ser thr gly)OPA ser val asp OPA arg ser arg AMB(pro glu pro
63/21   SEQ ID NO: 523   93/31   SEQ ID NO: 524
CCA TTC TGG TCG CAG GTC TTT CCG GAT GTT CAA GCA ACA AGT CGA CTA CAG GAA GCG GTG
pro phe trp ser gln val phe pro asp val gln ala thr ser arg leu gln glu ala val
123/41   153/51
AGA CCA CGA CCG CGG CAG GCA CGA CGG CAA GCC CCG GCG CCG CTC CGG GCC GAA GGT CGT
arg pro arg pro arg gln ala arg arg gln ala pro ala pro leu arg ala glu gly arg
183/61   213/71
CAT CGA CGG TAA GGA CCA GAA CGT CAC CGG CTC CGT GGT GTG CAC AAC CGC GGC CGG CAA
his arg arg)OCH(gly pro glu arg his arg leu arg gly val his asn arg gly arg gln
243/81   SEQ ID NO: 525   273/91
TGT CAA CAT CGC GAT CGG CGG GGC GGC GAC CGG CAT TGC CGC CGT GCT CAC CGA CGG CAA
cys gln his arg asp arg arg gly gly asp arg his cys arg arg ala his arg arg gln
303/101   333/111
CCC TCC GGA GGT GAA GTC CGT TGG GCT CGG TAA CGT CAA CGG CGT CAC GCT GGG ATA CAC
pro ser gly gly glu val arg trp ala arg)OCH(arg gln arg arg his ala gly ile his
363/121   393/131 SEQ ID NO: 526
GTC GGG CAC CGG ACA GGG TAA CGC CTC GGC AAC CAA GGA CGG CAG CCA CTA CAA GAT C
val gly his arg thr gly)OCH(arg leu gly asn gln gly arg gln pro leu gln asp)
                      SEQ ID NO: 527

SEQ ID NOS:522-527

FIG. 24C

Direct primer

SEQ ID NO: 528   5' ACG CGG CGC AGC CTG TTG 3'

SEQ ID No.528

FIG. 25

Reverse primer

SEQ ID NO: 529   5' CGA CCT TGG GAT TCG CCT 3'

SEQ ID No.529

FIG. 26

SEQ ID NO: 530   31/11
CCT ACC AGC AAG AGC CCA GGG CTT CAC AGG ACC TAA AAG GAG TAG CGC CCA TGG GCT TGA
(pro thr ser lys ser pro gly leu his arg thr)OCH lys glu AMB(arg pro trp ala)OPA
61/21   SEQ ID NO: 531   91/31   SEQ ID NO: 532
TCC AAT TTT CCT TCC GCC CCG TGC AAT ACC ATC TGC AAG ACC AGC GAC GGC CCG TGG TTG
(ser asn phe pro ser ala pro cys asn thr ile cys lys thr ser asp gly pro trp leu
121/41   SEQ ID NO: 533
CGG TCG CGC AGC TTG CGG AAA CGG GGT ATG GAC CCT GCC GTA CCG TTG TTG CCA CTT GAT
arg ser arg ser leu arg lys arg gly met asp pro ala val pro leu leu pro leu asp
181/61   211/71
GTC GTC GCT CTC CAC CCG TCG GGG GGC GAA AGC CAT TCC GAC ACT GGG ATC CTC AAA ACG
val val ala leu his pro ser gly gly glu ser his ser asp thr gly ile leu lys thr
241/81   271/91
TCG GCT GAG TGT CTG CAG GGC TCC GGG GAG CAG CCG ATC ATC ACC ATG TAC GAA CTG AAT
ser ala glu cys leu gln gly ser gly glu gln pro ile ile thr met tyr glu leu asn
301/101   331/111
AAG TCC CCC CCG CGC GAC TTC CAG ACA TTT GTT GTG GTT TCG GTT GAG GCC GAG GCG AGG
lys ser pro pro arg asp phe gln thr phe val val val ser val glu ala glu ala arg
361/121   391/131
CTC ATT TCG CAG CAA GCG GTC TCC GGG TCG CAG CAT CGT TGC GGC GAT CGC GGC GCA GTC
leu ile ser gln gln ala val ser gly ser gln his arg cys gly asp arg gly ala val
421/141
GTC GGA CGA GTC GTC GTC AAC GAC CAC GAT C
val gly arg val val val asn asp his asp)

SEQ ID NOS:530-533

FIG. 27A

SEQ ID NO: 534   31/11
CTA CCA GCA AGA GCC CAG GGC TTC ACA GGA CCT AAA AGG AGT AGC GCC CAT GGG CTT GAT
(leu pro ala arg ala gln gly phe thr gly pro lys arg ser ser ala his gly leu asp
61/21   SEQ ID NO: 535   91/31
CCA ATT TTC CTT CCG CCC CGT GCA ATA CCA TCT GCA AGA CCA GCG ACG GCC CGT GGT TGC
pro ile phe leu pro pro arg ala ile pro ser ala arg pro ala thr ala arg gly cys
121/41   151/51
GGT CGC GCA GCT TGC GGA AAC GGG GTA TGG ACC CTG CCG TAC CGT TGT TGC CAC TTG ATG
gly arg ala ala cys gly asn gly val trp thr leu pro tyr arg cys cys his leu met
181/61   211/71
TCG TCG CTC TCC ACC CGT CGG GGG GCG AAA GCC ATT CCG ACA CTG GGA TCC TCA AAA CGT
ser ser leu ser thr arg arg gly ala lys ala ile pro thr leu gly ser ser lys arg
241/81   271/91
CGG CTG AGT GTC TGC AGG GCT CCG GGG AGC AGC CGA TCA TCA CCA TGT ACG AAC TGA ATA
arg leu ser val cys arg ala pro gly ser ser arg ser ser pro cys thr asn)OPA(ile
301/101   331/111   SEQ ID NO: 536
AGT CCC CCC CGC GCG ACT TCC AGA CAT TTG TTG TGG TTT CGG TTG AGG CCG AGG CGA GGC
ser pro pro arg ala thr ser arg his leu leu trp phe arg leu arg pro arg arg gly
361/121   391/131
TCA TTT CGC AGC AAG CGG TCT CCG GGT CGC AGC ATC GTT GCG GCG ATC GCG GCG CAG TCG
ser phe arg ser lys arg ser pro gly arg ser ile val ala ala ile ala ala gln ser
421/141
TCG GAC GAG TCG TCG TCA ACG ACC ACG ATC
ser asp glu ser ser ser thr thr thr ile)

SEQ ID NOS:534-536

FIG. 27B

SEQ ID NO: 537   33/11
```
TAC CAG CAA GAG CCC AGG GCT TCA CAG GAC CTA AAA GGA GTA GCG CCC ATG GGC TTG ATC
(tyr gln gln glu pro arg ala ser gln asp leu lys gly val ala pro met gly leu ile
63/21         SEQ ID NO: 538                     93/31
CAA TTT TCC TTC CGC CCC GTG CAA TAC CAT CTG CAA GAC CAG CGA CGG CCC GTG GTT GCG
gln phe ser phe arg pro val gln tyr his leu gln asp gln arg arg pro val val ala
123/41                                           153/51
GTC GCG CAG CTT GCG GAA ACG GGG TAT GGA CCC TGC CGT ACC GTT GTT GCC ACT TGA TGT
val ala gln leu ala glu thr gly tyr gly pro cys arg thr val val ala thr)OPA(cys
183/61                                           213/71         SEQ ID NO: 539
CGT CGC TCT CCA CCC GTC GGG GGG CGA AAG CCA TTC CGA CAC TGG GAT CCT CAA AAC GTC
arg arg ser pro pro val gly gly arg lys pro phe arg his trp asp pro gln asn val
243/81                                           273/91
GGC TGA GTG TCT GCA GGG CTC CGG GGA GCA GCC GAT CAT CAC CAT GTA CGA ACT GAA TAA
gly)OPA(val ser ala gly leu arg gly ala ala asp his his his val arg thr glu)OCH
303/101        SEQ ID NO: 540                    333/111
GTC CCC CCC GCG CGA CTT CCA GAC ATT TGT TGT GGT TTC GGT TGA GGC CGA GGC GAG GCT
(val pro pro ala arg leu pro asp ile cys cys gly phe gly)OPA(gly arg gly glu ala
363/121        SEQ ID NO: 541                    393/131         SEQ ID NO: 542
CAT TTC GCA GCA AGC GGT CTC CGG GTC GCA GCA TCG TTG CGG CGA TCG CGG CGC AGT CGT
his phe ala ala ser gly leu arg val ala ala ser leu arg arg ser arg arg ser arg
423/141
CGG ACG AGT CGT CGT CAA CGA CCA CGA TC
arg thr ser arg arg gln arg pro arg)
```

SEQ ID NOS:537-541

FIG. 27C

SEQ ID NO: 543   MKTGTATTRRRLLAVLIALALPGAAVALLAEPSATGASDPCAASEVAR
                          TVGSVAKSMGDYLDSHPETNQVMTAVLQQQVGPGSVASLKAHFEANPK
                          VASDLHALSQPLTDLSTRCSLPISGLQAIGLMQAVQGARR

SEQ ID No.543

FIG. 28

SEQ ID NO: 544
```
GTGGGCAAGC AGCTAGCCGC GCTCGCCGCG CTGGTCGGTG CGTGCATGCT CGCAGCCGGA   60
TGCACCAACG TGGTCGACGG GACCGCCGTG GCTGCCGACA AATCCGGACC ACTGCATCAG   120
GATCCGATAC CGGTTTCAGC GCTTGAAGGG CTGCTTCTCG ACTTGAGCCA GATCAATGCC   180
GCGCTGGGTG CGACATCGAT GAAGGTGTGG TTCAACGCCA AGGCAATGTG GGACTGGAGC   240
AAGAGCGTGG CCGACAAGAA TTGCCTGGCT ATCGACGGTC CAGCACAGGA AAAGGTCTAT   300
GCCGGCACCG GGTGGACCGC TATGCCGCGG CAACGGCTGG ATGACAGCAT CGATGACTCC   360
AAGAAACGCG ACCACTACGC CATTCAAGCG GTCGTCGGCT TCCCGACCGC ACATGATGCC   420
GAGGAGTTCT ACAGCTCCTC GGTGCAAAGC TGGAGCAGCT GCTCGAACCG CCGGTTTGTC   480
GAAGTCACCC CCGGACAGGA CGACGCCGCC TGGACTGTGG CTGACGTTGT CAACGACAAC   540
GGCATGCTCA GTAGCTCGCA GGTTCAGGAA GGCGGCGACG GATGGACCTG CCAGCGTGCC   600
CTGACTGCGC GCAACAACGT CACTATCGAC ATTGTCACGT GCGCCTATAG CCAACCGGAT   660
TTGGTGGCGA TTGGCATCGC TAACCAAATC GCGGCCAAGG TTGCTAAGCA GTAG          714
```

SEQ ID No.544

FIG. 29

SEQ ID NO: 545

MGKQLAALAALVGACMLAAGCTNVVDGTAVAADKSGPLHQDPIPVFTSALEGLLLDLSQINAALGATS
MKVWFNAKAMWDWSKSVADKNCLAIDGPAQEKVYAGTGFTWTAMRGQRLDDSIDDSKKRDHYAIQAVV
GFPTAHDAEEFYSSSVQSWSSCSNRRFVEVTFTPGQDDAAWTVADVVNDNGMLSSSQVQEGGDGWTCQ
RALTARNNVTIDIVTCAYSQPDLVFTAIGIANQIAAKVAKQ

SEQ ID No.545

FIG. 30

1/1   SEQ ID NO: 546   31/11
AGG CGA ATA CCC GCG AGG GCA GCG CGA CGG CGG CCC TGC CGG CGC CGT GGC TGC TGA ACA
(arg arg ile pro ala arg ala ala arg arg arg pro cys arg arg arg gly cys)OPA(thr
61/21   SEQ ID NO: 547   91/31   SEQ ID NO: 548
ACA CAT CCC AGC CGC GCA CGC TTC CGG TAT GCG GCA GGA TAA ACG ACC CCA ACA GCA CGA
thr his pro ser arg ala arg phe arg tyr ala ala gly)OCH(thr thr pro thr ala arg
121/41   151/51   SEQ ID NO: 549
ACA CCA GGA TTG CGA CAA CCA AAG CCC TCG CGC CTG GCT CGA TTT CGC GCG CAA CGC GGC
thr pro gly leu arg gln pro lys pro ser arg leu ala arg phe arg ala gln arg gly
181/61   211/71
GTT CTG CCG CCT CGA TCT CAG CGC GGA GGG CGT CGA GAT C
val leu pro pro arg ser gln arg gly gly arg arg asp)

SEQ ID NOS:546-549

FIG. 31A

1/1   SEQ ID NO: 550   31/11
GGC GAA TAC CCG CGA GGG CAG CGC GAC GGC GGC CCT GCC GGC GCC GTG GCT GCT GAA CAA
(gly glu tyr pro arg gly gln arg asp gly gly pro ala gly ala val ala ala glu gln
61/21   SEQ ID NO: 551   91/31
CAC ATC CCA GCC GCG CAC GCT TCC GGT ATG CGG CAG GAT AAA CGA CCC CAA CAG CAC GAA
his ile pro ala ala his ala ser gly met arg gln asp lys arg pro gln gln his glu
121/41   151/51
CAC CAG GAT TGC GAC AAC CAA AGC CCT CGC GCC TGG CTC GAT TTC GCG CGC AAC GCG GCG
his gln asp cys asp asn gln ser pro arg ala trp leu asp phe ala arg asn ala ala
181/61   211/71
TTC TGC CGC CTC GAT CTC AGC GCG GAG GGC GTC GAG ATC
phe cys arg leu asp leu ser ala glu gly val glu ile)

SEQ ID NOS:550-551

FIG. 31B

```
1/1    ―SEQ ID NO: 552                    31/11
GCG AAT ACC CGC GAG GGC AGC GCG ACG GCG GCC CTG CCG GCG CCG TGG CTG CTG AAC AAC
(ala asn thr arg glu gly ser ala thr ala ala leu pro ala pro trp leu leu asn asn
61/21   ―SEQ ID NO: 553                   91/31
ACA TCC CAG CCG CGC ACG CTT CCG GTA TGC GGC AGG ATA AAC GAC CCC AAC AGC ACG AAC
thr ser gln pro arg thr leu pro val cys gly arg ile asn asp pro asn ser thr asn
121/41                                    151/51
ACC AGG ATT GCG ACA ACC AAA GCC CTC GCG CCT GGC TCG ATT TCG CGC GCA ACG CGG CGT
thr arg ile ala thr thr lys ala leu ala pro gly ser ile ser arg ala thr arg arg
181/61                                    211/71
TCT GCC GCC TCG ATC TCA GCG CGG AGG GCG TCG AGA TC
ser ala ala ser ile ser ala arg arg ala ser arg)
```

SEQ ID NOS:552-553

FIG. 31C

ORF according to Cole et al. (Nature 393:537-544) and containing seq31A
```
1/1    ―SEQ ID NO: 554                    31/11
taa acg acc cca aca gca cga aca cca gga ttg cga caa cca aag ccc tcg cgc ctg gct
OCH(thr thr pro thr ala arg thr pro gly leu arg gln pro lys pro ser arg leu ala
61/21   ―SEQ ID NO: 555                           91/31
cga ttt cgc gcg caa cgc ggc gtt ctg ccg cct cga tct cag cgc gga ggg cgt cga gat
arg phe arg ala gln arg gly val leu pro pro arg ser gln arg gly gly arg arg asp
121/41                                    151/51
ccc cgg cgt cgt gtt cgt ggc tca tca tct gca tcc tcc ggg ctt ggc cgc gct gac cgg
pro arg arg arg val arg gly ser ser ser ala ser ser gly leu gly arg ala asp arg
181/61                                    211/71
cag ccc gac ccc agg cat gcc cag gcc gac ggc gcg ccc cgg ctg ccc ggc ggt gtg cgc
gln pro asp pro arg his ala gln ala asp gly ala pro arg leu pro gly gly val arg
241/81                                    271/91
gtc gcc ggc gcg ggt gcg gcg gtg ggt cag gac gcc ggc gtc ggc gat gag gtg gtg cgg
val ala gly ala gly ala ala val gly gln asp ala gly val gly asp glu val val arg
301/101                                   331/111
cgc cgc ttc ggt gac ctt cgt ggt gat gac gtc gcc ggg acg cac gcg cgg ctg gcc ggc
arg arg phe gly asp leu arg gly asp asp val ala gly thr his ala arg leu ala gly
361/121                                   391/131
ggt gaa gtg cac cag gcg ccc gtc gcg cgc ccg ccc gct cat gcg cgc cgt gac ggt gtc
gly glu val his gln ala pro val ala arg pro pro ala his ala arg arg asp gly val
421/141                                   451/151
ctt gcg ccc ttc ccc ggt ggc cac cag cac ctc gac ggc ctg ccc gac cag ggc gcg gtt
leu ala pro phe pro gly gly his gln his leu asp gly leu pro asp gln gly ala val
481/161                                   511/171
ggc ttc cag cga gat ttg ctc ctg cag cgc gat cag gcg ttc ata gcg ttc ctg cac aac
gly phe gln arg asp leu leu leu gln arg asp gln ala phe ile ala phe leu his asn
541/181                                   571/191
ggc ttt cgg cag ctg tcc gtc gag ttg cgc ggc cgg tgt ccc ggg ccg ctt gga gta ttg
gly phe arg gln leu ser val glu leu arg gly arg cys pro gly pro leu gly val leu
601/201                                   631/211
gaa ggt aaa tgc ggc cgc gaa gcg ggc ccg gcg cac cac gtc gag cgt ggc cgc gaa gtc
glu gly lys cys gly arg glu ala gly pro ala his his val glu arg gly arg glu val
661/221                                   691/231
ctc ttc ggt ctc ccc ggg gaa acc gac gat cag atc ggt ggt aat cgc ggc atg cgg gat
leu phe gly leu pro gly glu thr asp asp gln ile gly gly asn arg gly met arg asp
721/241                                   751/251
ggc cgc ccg cac gcg ctc gat gat gcc gag gta gcg ctc ggc acg ata gga ccg ccg cat
gly arg pro his ala leu asp asp ala glu val ala leu gly thr ile gly pro pro his
781/261                                   811/271
cgc gcg cag gat ccg gtc gga tcc gga ctg tag
arg ala gln asp pro val gly ser gly leu)AMB
```

SEQ ID NOS:554-555

FIG. 31D

```
1/1 ──SEQ ID NO: 556                        31/11
aga ctg gtg tac acg gag acc aag ctg aac     tcg gca ttc tcc ttc ggc ggg cct aag tgt
(arg leu val tyr thr glu thr lys leu asn    ser ala phe ser phe gly gly pro lys cys
61/21 ─SEQ ID NO: 557                       91/31
cta gtg aag gtc att cag aaa ctg tcg ggc     ttg agc atc aac cgg ttc atc gcg att gac
leu val lys val ile gln lys leu ser gly     leu ser ile asn arg phe ile ala ile asp
121/41                                      151/51
ttc gtc ggt ttc gcg cgg atg gtc gag gcc     ctc ggc ggc gtc gag gta tgc agc acc acc
phe val gly phe ala arg met val glu ala     leu gly gly val glu val cys ser thr thr
181/61                                      211/71
ccg ttg cgg gac tac gaa ctg ggc acg gtg     ctg gag cac gcc gga cgc cag gtc att gac
pro leu arg asp tyr glu leu gly thr val     leu glu his ala gly arg gln val ile asp
241/81                                      271/91
ggg ccg acc gcg ctg aac tat gtg cgc gct     cgc cag gtc acc acc gag agc aat ggc gac
gly pro thr ala leu asn tyr val arg ala     arg gln val thr thr glu ser asn gly asp
301/101                                     331/111
tac ggg cgc atc aaa cgc cag cag ttg ttt     ttg tcg tcg ctg ctg cgt tcg atg atc
tyr gly arg ile lys arg gln gln leu phe     leu ser ser leu leu arg ser met ile)
```

SEQ ID NOS:556-557

FIG. 32A

```
1/1 ──SEQ ID NO: 558                        31/11
gac tgg tgt aca cgg aga cca agc tga act     cgg cat tct cct tcg gcg ggc cta agt gtc
(asp trp cys thr arg arg pro ser)OPA(thr    arg his ser pro ser ala gly leu ser val)
61/21 ─SEQ ID NO: 559  SEQ ID NO: 560       91/31
tag tga agg tca ttc aga aac tgt cgg gct     tga gca tca acc ggt tca tcg cga ttg act
AMB OPA(arg ser phe arg asn cys arg ala)OPA(ala ser thr gly ser ser arg leu thr
121/41    ──SEQ ID NO: 561                  151/51   ──SEQ ID NO: 562
tcg tcg gtt tcg cgc gga tgg tcg agg ccc     tcg gcg gcg tcg agg tat gca gca cca ccc
ser ser val ser arg gly trp ser arg pro     ser ala ala ser arg tyr ala ala pro pro
181/61                                      211/71
cgt tgc ggg act acg aac tgg gca cgg tgc     tgg agc acg ccg gac gcc agg tca ttg acg
arg cys gly thr thr asn trp ala arg cys     trp ser thr pro asp ala arg ser leu thr
241/81                                      271/91
ggc cga ccg cgc tga act atg tgc gcg ctc     gcc agg tca cca ccg aga gca atg gcg act
gly arg pro arg)OPA(thr met cys ala leu     ala arg ser pro pro arg ala met ala thr
301/101          ──SEQ ID NO: 563           331/111
acg ggc gca tca aac gcc agc agt tgt ttt     tgt cgt cgc tgc tgc gtt cga tga tc
thr gly ala ser asn ala ser ser cys phe     cys arg arg cys cys val arg)OPA
```

SEQ ID NOS:558-563

FIG. 32B

1/1 _SEQ ID NO:564   31/11
act ggt gta cac gga gac caa gct gaa ctc ggc att ctc ctt cgg cgg gcc taa gtg tct
(thr gly val his gly asp gln ala glu leu gly ile leu leu arg arg ala)OCH(val ser
61/21 _SEQ ID NO: 565   91/31   SEQ ID NO: 566
agt gaa ggt cat tca gaa act gtc ggg ctt gag cat caa ccg gtt cat cgc gat tga ctt
ser glu gly his ser glu thr val gly leu glu his gln pro val his arg asp)OP(leu
121/41   151/51   SEQ ID NO: 567
cgt cgg ttt cgc gcg gat ggt cga ggc cct cgg cgg cgt cga ggt atg cag cac cac ccc
arg arg phe arg ala asp gly arg gly pro arg arg arg arg gly met gln his his pro
181/61   211/71
gtt gcg gga cta cga act ggg cac ggt gct gga gca cgc cgg acg cca ggt cat tga cgg
val ala gly leu arg thr gly his gly ala gly ala arg arg thr pro gly his)OPA(arg
241/81   271/91   SEQ ID NO: 568
gcc gac cgc gct gaa cta tgt gcg cgc tcg cca ggt cac cac cga gag caa tgg cga cta
ala asp arg ala glu leu cys ala arg ser pro gly his his arg glu gln trp arg leu
301/101   331/111
cgg gcg cat caa acg cca gca gtt gtt ttt gtc gtc gct gct gcg ttc gat gat c
arg ala his gln thr pro ala val val phe val val ala ala ala phe asp asp)

SEQ ID NOS:564-568

FIG. 32C

sequence Rv0822c predicted by Cole et al. (Nature 393:537-544) and containing
seq 32A

1/1 _SEQ ID NO: 569   31/11
atg agt gac ggc gag agc gcc gcg ccg tgg gca cgg ctc tcc gag tca gca ttc ccc gat
(Met ser asp gly glu ser ala ala pro trp ala arg leu ser glu ser ala phe pro asp
61/21 _SEQ ID NO: 570   91/31
ggt gtt gac cga tgg atc acg gta ccg ccc gcc aca tgg gtg gca gcc cag ggt ccg cgg
gly val asp arg trp ile thr val pro pro ala thr trp val ala ala gln gly pro arg
121/41   151/51
gac acc cag aat gtc ggc tgt cat gcc acc ggc gcc gtt agt gtg gcc gat ctg atc gcc
asp thr gln asn val gly cys his ala thr gly ala val ser val ala asp leu ile ala
181/61   211/71
agg ctc ggc ccc gct ttt cct gac ctc ccc acg cac cgc cat gtc gcc ccc gaa ccc gag
arg leu gly pro ala phe pro asp leu pro thr his arg his val ala pro glu pro glu
241/81   271/91
cca tcc ggc cgc ggc ccg aag gtc cac gac gac gcc gac gac cag cag gac acc gag gct
pro ser gly arg gly pro lys val his asp asp ala asp asp gln gln asp thr glu ala
301/101   331/111
atc gcc atc ccg gcc cac tcg ctc gag ttc ctc tcg gag ctt ccc gac ctc cgg gca gcc
ile ala ile pro ala his ser leu glu phe leu ser glu leu pro asp leu arg ala ala
361/121   391/131
aac tat ccg cgc gcc gac cac gcc cgc cgt gaa ccc gag cta ccc ggc aag cag cta acc
asn tyr pro arg ala asp his ala arg arg glu pro glu leu pro gly lys gln leu thr
421/141   451/151
gga tcg gct cga gtg cgg cca ttg cgg atc cgc cga acg tcg ccc gcg ccc gcc aag cca
gly ser ala arg val arg pro leu arg ile arg arg thr ser pro ala pro ala lys pro
481/161   511/171
gcg ccg aac tcc ggc cgg cgc ccg atg gtg ctg gcc gcg cgc tcg ctg gcg gct ctg ttt
ala pro asn ser gly arg arg pro met val leu ala ala arg ser leu ala ala leu phe
541/181   571/191
gcc gct ctg gcg ttg gcg ctg acc ggc ggg gca tgg cag tgg agc gcg tcg aag aac agc
ala ala leu ala leu ala leu thr gly gly ala trp gln trp ser ala ser lys asn ser
601/201   631/211
cgg ctg aac atg gta agc gcg ctc gac ccg cat tcg ggc gac atc gtc aac ccc agc ggg
arg leu asn met val ser ala leu asp pro his ser gly asp ile val asn pro ser gly

SEQ ID NOS:569-570

FIG. 32D

661/221
cag cat ggc gac gag aac ttc ttg ctc gtc
gln his gly asp glu asn phe leu leu val
721/241
aat atc ggc gcc ggc gac gcc gag gac gcc
asn ile gly ala gly asp ala glu asp ala
781/261
gtc aac att ccg gcc agc cgc gag cgg gtc
val asn ile pro ala ser arg glu arg val
841/281
atc act cca atc caa tgc gag gcg tgg aac
ile thr pro ile gln cys glu ala trp asn
901/301
gac gag aag acg gga acg atg ggt ccc aga
asp glu lys thr gly thr met gly pro arg
961/321
gca ttc tcc ttc ggc ggg cct aag tgt cta
ala phe ser phe gly gly pro lys cys leu
1021/341
agc atc aac cgg ttc atc gcg att gac ttc
ser ile asn arg phe ile ala ile asp phe
1081/361
ggc ggc gtc gag gta tgc agc acc acc ccg
gly gly val glu val cys ser thr thr pro
1141/381
gag cac gcc gga cgc cag gtc att gac ggg
glu his ala gly arg gln val ile asp gly
1201/401
cag gtc acc acc gag agc aat ggc gac tac
gln val thr thr glu ser asn gly asp tyr
1261/421
tcg tcg ctg ctg cgt tcg atg atc tcg acg
ser ser leu leu arg ser met ile ser thr
1321/441
aac gtc gtc aac atg ttc atc ggt aac agc
asn val val asn met phe ile gly asn ser
1381/461
gtc gaa ctc ggt cga tcg ttg cag cat atg
val glu leu gly arg ser leu gln his met
1441/481
ccg acc ggt ata acc gac cag aac ggc gac
pro thr gly ile thr asp gln asn gly asp
1501/501
ctt ttc acc gcc atc atc gac gac gat ccg
leu phe thr ala ile ile asp asp asp pro
1561/521
cgt ctg ggc aac acg ccg tcg acc ccg ccg
arg leu gly asn thr pro ser thr pro pro
1621/541
ctg acc aac gag att cag cac cag cag gtt
leu thr asn glu ile gln his gln gln val
1681/561
cag gtc tct aac tcg acc ggc cag gcc ggt
gln val ser asn ser thr gly gln ala gly
1741/581
cgg aac ggc ttc aac gtg atg gct ccg gac
arg asn gly phe asn val met ala pro asp
1801/601
gtg ttt ttt tcg ccc ggc aac gaa cag gct
val phe phe ser pro gly asn glu gln ala
1861/621
tca aag atc gag cgg gtg acc ggg atc ggc
ser lys ile glu arg val thr gly ile gly

691/231
ggt atg gac tct cgt gcc ggg gcg aac gcc
gly met asp ser arg ala gly ala asn ala
751/251
ggc ggc gca cgt tcg gac acc gtc atg ctg
gly gly ala arg ser asp thr val met leu
811/271
gtc gcg gtg tcg ttc ccc cgc gac ctg gcg
val ala val ser phe pro arg asp leu ala
871/291
ccc gag acc ggt aag tac gga ccc atc tac
pro glu thr gly lys tyr gly pro ile tyr
931/311
ctg gtg tac acg gag acc aag ctg aac tcg
leu val tyr thr glu thr lys leu asn ser
991/331
gtg aag gtc att cag aaa ctg tcg ggc ttg
val lys val ile gln lys leu ser gly leu
1051/351
gtc ggt ttc gcg cgg atg gtc gag gcc ctc
val gly phe ala arg met val glu ala leu
1111/371
ttg cgg gac tac gaa ctg ggc acg gtg ctg
leu arg asp tyr glu leu gly thr val leu
1171/391
ccg acc gcg ctg aac tat gtg cgc gct cgc
pro thr ala leu asn tyr val arg ala arg
1231/411
ggg cgc atc aaa cgc cag cag ttg ttt ttg
gly arg ile lys arg gln gln leu phe leu
1291/431
gac acc ttg ttc aac ctc agc agg ctc aac
asp thr leu phe asn leu ser arg leu asn
1351/451
tac gtg gac aac gtc aag acc aaa gac ctg
tyr val asp asn val lys thr lys asp leu
1411/471
gcg gcc ggg cac gtc acg ttc gtg acc gtt
ala ala gly his val thr phe val thr val
1471/491
gag ccc ccg cgt acc tcc gac atg aag gcg
glu pro pro arg thr ser asp met lys ala
1531/511
ctg ccc ctg gaa aac gat cac aac gcc cag
leu pro leu glu asn asp his asn ala gln
1591/531
acc acc acc aag aag gcg ccg cag gcg ggt
thr thr thr lys lys ala pro gln ala gly
1651/551
acg acg acc tcg cca aaa gag gtc aca gtg
thr thr thr ser pro lys glu val thr val
1711/571
ttg gcc acc acc gcc acc gat cag ctc aag
leu ala thr thr ala thr asp gln leu lys
1771/591
gac tac ccg agt tcg ctg ctg gcc acc aca
asp tyr pro ser ser leu leu ala thr thr
1831/611
gcc gcc acc gtg gcc gcc gtg ttc ggc cag
ala ala thr val ala ala val phe gly gln
1891/631
caa ctg gtc cag gtg gtg ctg ggc caa gac
gln leu val gln val val leu gly gln asp

SEQ ID NOS:569-570 (continued 1)

FIG. 32D (continued 1)

```
1921/641                              1951/651
ttc agc gcg gtg cgc gct ccc ctg ccg agt ggc tcc acc gtc agc gtg cag ata agc cgc
phe ser ala val arg ala pro leu pro ser gly ser thr val ser val gln ile ser arg
1981/661                              2011/671
aac tcc tcc agc cca ccg acc aag ctg ccc gag gac ctg acg gtc acc aac gcc gcc gac
asn ser ser ser pro pro thr lys leu pro glu asp leu thr val thr asn ala ala asp
2041/681
acc acc tgc gag tag
thr thr cys glu)AMB
```

SEQ ID NOS:569-570 (continued 2)
FIG. 32D (continued 2)

ORF according to Cole et al. (Nature 393:537-544) and containing Rv0822c

```
1/1    SEQ ID NO: 571               31/11
tag gac atg agt gac ggc gag agc gcc gcg ccg tgg gca cgg ctc tcc gag tca gca ttc
AMB(asp met ser asp gly glu ser ala ala pro trp ala arg leu ser glu ser ala phe
61/21   SEQ ID NO: 572              91/31
ccc gat ggt gtt gac cga tgg atc acg gta ccg ccc gcc aca tgg gtg gca gcc cag ggt
pro asp gly val asp arg trp ile thr val pro pro ala thr trp val ala ala gln gly
121/41                              151/51
ccg cgg gac acc cag aat gtc ggc tgt cat gcc acc ggc gcc gtt agt gtg gcc gat ctg
pro arg asp thr gln asn val gly cys his ala thr gly ala val ser val ala asp leu
181/61                              211/71
atc gcc agg ctc ggc ccc gct ttt cct gac ctc ccc acg cac cgc cat gtc gcc ccc gaa
ile ala arg leu gly pro ala phe pro asp leu pro thr his arg his val ala pro glu
241/81                              271/91
ccc gag cca tcc ggc cgc ggc ccg aag gtc cac gac gac gcc gac gac cag cag gac acc
pro glu pro ser gly arg gly pro lys val his asp asp ala asp asp gln gln asp thr
301/101                             331/111
gag gct atc gcc atc ccg gcc cac tcg ctc gag ttc ctc tcg gag ctt ccc gac ctc cgg
glu ala ile ala ile pro ala his ser leu glu phe leu ser glu leu pro asp leu arg
361/121                             391/131
gca gcc aac tat ccg cgc gcc gac cac gcc cgc cgt gaa ccc gag cta ccc ggc aag cag
ala ala asn tyr pro arg ala asp his ala arg arg glu pro glu leu pro gly lys gln
421/141                             451/151
cta acc gga tcg gct cga gtg cgg cca ttg cgg atc cgc cga acg tcg ccc gcg ccc gcc
leu thr gly ser ala arg val arg pro leu arg ile arg arg thr ser pro ala pro ala
481/161                             511/171
aag cca gcg ccg aac tcc ggc cgg cgc ccg atg gtg ctg gcc gcg cgc tcg ctg gcg gct
lys pro ala pro asn ser gly arg arg pro met val leu ala ala arg ser leu ala ala
541/181                             571/191
ctg ttt gcc gct ctg gcg ttg gcg ctg acc ggc ggg gca tgg cag tgg agc gcg tcg aag
leu phe ala ala leu ala leu ala leu thr gly gly ala trp gln trp ser ala ser lys
601/201                             631/211
aac agc cgg ctg aac atg gta agc gcg ctc gac ccg cat tcg ggc gac atc gtc aac ccc
asn ser arg leu asn met val ser ala leu asp pro his ser gly asp ile val asn pro
661/221                             691/231
agc ggg cag cat ggc gac gag aac ttc ttg ctc gtc ggt atg gac tct cgt gcc ggg gcg
ser gly gln his gly asp glu asn phe leu leu val gly met asp ser arg ala gly ala
721/241                             751/251
aac gcc aat atc ggc gcc ggc gac gcc gag gac gcc ggc ggc gca cgt tcg gac acc gtc
asn ala asn ile gly ala gly asp ala glu asp ala gly gly ala arg ser asp thr val
781/261                             811/271
atg ctg gtc aac att ccg gcc agc cgc gag cgg gtc gtc gcg gtg tcg ttc ccc cgc gac
met leu val asn ile pro ala ser arg glu arg val val ala val ser phe pro arg asp
```

SEQ ID NOS: 571-572

FIG. 32E

```
841/281                              871/291
ctg gcg atc act cca atc caa tgc gag gcg  tgg aac ccc gag acc ggt aag tac gga ccc
leu ala ile thr pro ile gln cys glu ala  trp asn pro glu thr gly lys tyr gly pro
901/301                              931/311
atc tac gac gag aag acg gga acg atg ggt  ccc aga ctg gtg tac acg gag acc aag ctg
ile tyr asp glu lys thr gly thr met gly  pro arg leu val tyr thr glu thr lys leu
961/321                              991/331
aac tcg gca ttc tcc ttc ggc ggg cct aag  tgt cta gtg aag gtc att cag aaa ctg tcg
asn ser ala phe ser phe gly gly pro lys  cys leu val lys val ile gln lys leu ser
1021/341                             1051/351
ggc ttg agc atc aac cgg ttc atc gcg att  gac ttc gtc ggt ttc gcg cgg atg gtc gag
gly leu ser ile asn arg phe ile ala ile  asp phe val gly phe ala arg met val glu
1081/361                             1111/371
gcc ctc ggc ggc gtc gag gta tgc agc acc  acc ccg ttg cgg gac tac gaa ctg ggc acg
ala leu gly gly val glu val cys ser thr  thr pro leu arg asp tyr glu leu gly thr
1141/381                             1171/391
gtg ctg gag cac gcc gga cgc cag gtc att  gac ggg ccg acc gcg ctg aac tat gtg cgc
val leu glu his ala gly arg gln val ile  asp gly pro thr ala leu asn tyr val arg
1201/401                             1231/411
gct cgc cag gtc acc acc gag agc aat ggc  gac tac ggg cgc atc aaa cgc cag cag ttg
ala arg gln val thr thr glu ser asn gly  asp tyr gly arg ile lys arg gln gln leu
1261/421                             1291/431
ttt ttg tcg tcg ctg ctg cgt tcg atg atc  tcg acg gac acc ttg ttc aac ctc agc agg
phe leu ser ser leu leu arg ser met ile  ser thr asp thr leu phe asn leu ser arg
1321/441                             1351/451
ctc aac aac gtc gtc aac atg ttc atc ggt  aac agc tac gtg gac aac gtc aag acc aaa
leu asn asn val val asn met phe ile gly  asn ser tyr val asp asn val lys thr lys
1381/461                             1411/471
gac ctg gtc gaa ctc ggt cga tcg ttg cag  cat atg gcg gcc ggg cac gtc acg ttc gtg
asp leu val glu leu gly arg ser leu gln  his met ala ala gly his val thr phe val
1441/481                             1471/491
acc gtt ccg acc ggt ata acc gac cag aac  ggc gac gag ccc ccg cgt acc tcc gac atg
thr val pro thr gly ile thr asp gln asn  gly asp glu pro pro arg thr ser asp met
1501/501                             1531/511
aag gcg ctt ttc acc gcc atc atc gac gac  gat ccg ctg ccc ctg gaa aac gat cac aac
lys ala leu phe thr ala ile ile asp asp  asp pro leu pro leu glu asn asp his asn
1561/521                             1591/531
gcc cag cgt ctg ggc aac acg ccg tcg acc  ccg ccg acc acc acc aag aag gcg ccg cag
ala gln arg leu gly asn thr pro ser thr  pro pro thr thr thr lys lys ala pro gln
1621/541                             1651/551
gcg ggt ctg acc aac gag att cag cac cag  cag gtt acg acg acc tcg cca aaa gag gtc
ala gly leu thr asn glu ile gln his gln  gln val thr thr thr ser pro lys glu val
1681/561                             1711/571
aca gtg cag gtc tct aac tcg acc ggc cag  gcc ggt ttg gcc acc acc gcc acc gat cag
thr val gln val ser asn ser thr gly gln  ala gly leu ala thr thr ala thr asp gln
1741/581                             1771/591
ctc aag cgg aac ggc ttc aac gtg atg gct  ccg gac gac tac ccg agt tcg ctg ctg gcc
leu lys arg asn gly phe asn val met ala  pro asp asp tyr pro ser ser leu leu ala
1801/601                             1831/611
acc aca gtg ttt ttt tcg ccc ggc aac gaa  cag gct gcc gcc acc gtg gcc gcc gtg ttc
thr thr val phe phe ser pro gly asn glu  gln ala ala ala thr val ala ala val phe
1861/621                             1891/631
ggc cag tca aag atc gag cgg gtg acc ggg  atc ggc caa ctg gtc cag gtg gtg ctg ggc
gly gln ser lys ile glu arg val thr gly  ile gly gln leu val gln val val leu gly
1921/641                             1951/651
caa gac ttc agc gcg gtg cgc gct ccc ctg  ccg agt ggc tcc acc gtc agc gtg cag ata
gln asp phe ser ala val arg ala pro leu  pro ser gly ser thr val ser val gln ile
1981/661                             2011/671
agc cgc aac tcc tcc agc cca ccg acc aag  ctg ccc gag gac ctg acg gtc acc aac gcc
ser arg asn ser ser ser pro pro thr lys  leu pro glu asp leu thr val thr asn ala
2041/681
gcc gac acc acc tgc gag tag
ala asp thr thr cys glu)AMB
```

SEQ ID NOS: 571-572 (continued 1)

FIG. 32E (continued 1)

```
1/1  ⌐SEQ ID NO: 573                          31/11
CGT CAC CTC TGC CAT GGT CCA TCT ACG GTA   TCT GCG ACA AGG GCA GCG TCG ATC CCT CGA
(arg his leu cys his gly pro ser thr val   ser ala thr arg ala ala ser ile pro arg
61/21  ⌐SEQ ID NO: 574                       91/31
CAT GCA GAG TCG GTG TTC GCT TCA CGC GAA   CTA GGC GCG CCT AGC CTG GAC GAG TCC CCG
his ala glu ser val phe ala ser arg glu   leu gly ala pro ser leu asp glu ser pro
121/41                                       151/51
GGC CGA CAT TCG CCC GAG GCC TTG GCC TCC   ATC ACC TAA TTG TGT GCA AAA CCG TAT CTA
gly arg his ser pro glu ala leu ala ser   ile thr)OCH(leu cys ala lys pro tyr leu
181/61                                       211/71          ⌐SEQ ID NO: 575
ATT GAT ACG ATT GCG CAC ATG GCT ATC TGG   GAT C
ile asp thr ile ala his met ala ile trp   asp)
```

SEQ ID NOS:573-575

FIG. 33A

```
1/1  ⌐SEQ ID NO: 576                          31/11
GTC ACC TCT GCC ATG GTC CAT CTA CGG TAT   CTG CGA CAA GGG CAG CGT CGA TCC CTC GAC
(val thr ser ala met val his leu arg tyr   leu arg gln gly gln arg arg ser leu asp
61/21  ⌐SEQ ID NO: 577                       91/31
ATG CAG AGT CGG TGT TCG CTT CAC GCG AAC   TAG GCG CGC CTA GCC TGG ACG AGT CCC CGG
met gln ser arg cys ser leu his ala asn)   AMB(ala arg leu ala trp thr ser pro arg
121/41                                       151/51  ⌐SEQ ID NO: 578
GCC GAC ATT CGC CCG AGG CCT TGG CCT CCA   TCA CCT AAT TGT GTG CAA AAC CGT ATC TAA
ala asp ile arg pro arg pro trp pro pro   ser pro asn cys val gln asn arg ile)OCH
181/61                                       211/71
TTG ATA CGA TTG CGC ACA TGG CTA TCT GGG   ATC
(leu ile arg leu arg thr trp leu ser gly   ile)
 ⌐SEQ ID NO: 579
```

SEQ ID NOS:576-579

FIG. 33B

```
1/1  ⌐SEQ ID NO: 580                          31/11
CCG TCA CCT CTG CCA TGG TCC ATC TAC GGT   ATC TGC GAC AAG GGC AGC GTC GAT CCC TCG
(pro ser pro leu pro trp ser ile tyr gly   ile cys asp lys gly ser val asp pro ser
61/21  ⌐SEQ ID NO: 581                       91/31
ACA TGC AGA GTC GGT GTT CGC TTC ACG CGA   ACT AGG CGC GCC TAG CCT GGA CGA GTC CCC
thr cys arg val gly val arg phe thr arg   thr arg arg ala)AMB(pro gly arg val pro
121/41                                       151/51          ⌐SEQ ID NO: 582
GGG CCG ACA TTC GCC CGA GGC CTT GGC CTC   CAT CAC CTA ATT GTG TGC AAA ACC GTA TCT
gly pro thr phe ala arg gly leu gly leu   his his leu ile val cys lys thr val ser
181/61                                       211/71
AAT TGA TAC GAT TGC GCA CAT GGC TAT CTG   GGA TC
asn)OPA(tyr asp cys ala his gly tyr leu   gly)
 ⌐SEQ ID NO: 583
```

SEQ ID NOS:580-583

FIG. 33C

sequence Rv1044 predicted by Cole et al. (Nature 393:537-544) and
containing seq33A

1/1 SEQ ID NO: 584   31/11
ttg tgt gca aaa ccg tat cta att gat acg att gcg cac atg gct atc tgg gat cgc ctc
(leu cys ala lys pro tyr leu ile asp thr ile ala his met ala ile trp asp arg leu
61/21 SEQ ID NO: 585   91/31
gtc gag gtt gcc gcc gag caa cat ggc tac gtc acg act cgc gat gcg cga gac atc ggc
val glu val ala ala glu gln his gly tyr val thr thr arg asp ala arg asp ile gly
121/41   151/51
gtc gac cct gtg cag ctc cgc ctc cta gcg ggg cgc gga cgt ctt gag cgt gtc ggc cga
val asp pro val gln leu arg leu leu ala gly arg gly arg leu glu arg val gly arg
181/61   211/71
ggt gtg tac cgg gtg ccc gtg ctg ccg cgt ggt gag cac gac gat ctc gca gcc gca gtg
gly val tyr arg val pro val leu pro arg gly glu his asp asp leu ala ala ala val
241/81   271/91
tcg tgg act ttg ggg cgt ggc gtt atc tcg cat gag tcg gcc ttg gcg ctt cat gcc ctc
ser trp thr leu gly arg gly val ile ser his glu ser ala leu ala leu his ala leu
301/101   331/111
gct gac gtg aac ccg tcg cgc atc cat ctc acc gtc ccg cgc aac aac cat ccg cgt gcg
ala asp val asn pro ser arg ile his leu thr val pro arg asn asn his pro arg ala
361/121   391/131
gcc ggg ggc gag ctg tac cga gtt cac cgc cgc gac ctc cag gca gcc cac gtc act tcg
ala gly gly glu leu tyr arg val his arg arg asp leu gln ala ala his val thr ser
421/141   451/151
gtc gac gga ata ccc gtc acg acg gtt gcg cgc acc atc aaa gac tgc gtg aag acg ggc
val asp gly ile pro val thr thr val ala arg thr ile lys asp cys val lys thr gly
481/161   511/171
acg gat cct tat cag ctt cgg gcc gcg atc gag cga gcc gaa gcc gag ggc acg ctt cgt
thr asp pro tyr gln leu arg ala ala ile glu arg ala glu ala glu gly thr leu arg
541/181   571/191
cgt ggg tca gca gct gag cta cgc gct gcg ctc gat gag acc act gcc gga tta cgc gct
arg gly ser ala ala glu leu arg ala ala leu asp glu thr thr ala gly leu arg ala
601/201
cgg ccg aag cga gca tcg gcg tga
arg pro lys arg ala ser ala)OPA

SEQ ID NOS:584-585

FIG. 33D

ORF according to Cole et al. (Nature 393:537-544) and containing Rv1044

1/1 SEQ ID NO: 586   31/11
taa ttg tgt gca aaa ccg tat cta att gat acg att gcg cac atg gct atc tgg gat cgc
OCH(leu cys ala lys pro tyr leu ile asp thr ile ala his met ala ile trp asp arg
61/21 SEQ ID NO: 587   91/31
ctc gtc gag gtt gcc gcc gag caa cat ggc tac gtc acg act cgc gat gcg cga gac atc
leu val glu val ala ala glu gln his gly tyr val thr thr arg asp ala arg asp ile
121/41   151/51
ggc gtc gac cct gtg cag ctc cgc ctc cta gcg ggg cgc gga cgt ctt gag cgt gtc ggc
gly val asp pro val gln leu arg leu leu ala gly arg gly arg leu glu arg val gly
181/61   211/71
cga ggt gtg tac cgg gtg ccc gtg ctg ccg cgt ggt gag cac gac gat ctc gca gcc gca
arg gly val tyr arg val pro val leu pro arg gly glu his asp asp leu ala ala ala
241/81   271/91
gtg tcg tgg act ttg ggg cgt ggc gtt atc tcg cat gag tcg gcc ttg gcg ctt cat gcc
val ser trp thr leu gly arg gly val ile ser his glu ser ala leu ala leu his ala
301/101   331/111
ctc gct gac gtg aac ccg tcg cgc atc cat ctc acc gtc ccg cgc aac aac cat ccg cgt
leu ala asp val asn pro ser arg ile his leu thr val pro arg asn asn his pro arg
361/121   391/131
gcg gcc ggg ggc gag ctg tac cga gtt cac cgc cgc gac ctc cag gca gcc cac gtc act
ala ala gly gly glu leu tyr arg val his arg arg asp leu gln ala ala his val thr
421/141   451/151
tcg gtc gac gga ata ccc gtc acg acg gtt gcg cgc acc atc aaa gac tgc gtg aag acg
ser val asp gly ile pro val thr thr val ala arg thr ile lys asp cys val lys thr
481/161   511/171
ggc acg gat cct tat cag ctt cgg gcc gcg atc gag cga gcc gaa gcc gag ggc acg ctt
gly thr asp pro tyr gln leu arg ala ala ile glu arg ala glu ala glu gly thr leu
541/181   571/191
cgt cgt ggg tca gca gct gag cta cgc gct gcg ctc gat gag acc act gcc gga tta cgc
arg arg gly ser ala ala glu leu arg ala ala leu asp glu thr thr ala gly leu arg
601/201
gct cgg ccg aag cga gca tcg gcg tga
ala arg pro lys arg ala ser ala)OPA

SEQ ID NOS:586-587

FIG. 33E

1/1 SEQ ID NO: 588   31/11
ATC CAA CCT GCT GGG CCT GCG CCT TCG AAT CGA CGG CCA GGC CAC CGC TCG CTG CCG GCA
(ile gln pro ala gly pro ala pro ser asn arg arg pro gly his arg ser leu pro ala
61/21 SEQ ID NO: 589   91/31
ACA ACA CCT GGA ATG GGG ACC TTT TCG GTG TTG CTG GTA ACC GGG ACA ACC GGC ACC ACG
thr thr pro gly met gly thr phe ser val leu leu val thr gly thr thr gly thr thr
121/41   151/51
CCT CGG TCG AGA CGT ATC GCG GCA GCG TTG GCC CTG TCG TTG CTG ACA ATT ACC GCT GGC
pro arg ser arg arg ile ala ala ala leu ala leu ser leu leu thr ile thr ala gly
181/61   211/71
CGC CGC ATA TTT GCC GCG CTG CCG CGG GCC GGA TC
arg arg ile phe ala ala leu pro arg ala gly)

SEQ ID NOS:588-589

FIG. 34A

1/1 ⟋SEQ ID NO: 590   31/11
TCC AAC CTG CTG GGC CTG CGC CTT CGA ATC GAC GGC CAG GCC ACC GCT CGC TGC CGG CAA
(ser asn leu leu gly leu arg leu arg ile asp gly gln ala thr ala arg cys arg gln
61/21 ⟋SEQ ID NO: 591   91/31
CAA CAC CTG GAA TGG GGA CCT TTT CGG TGT TGC TGG TAA CCG GGA CAA CCG GCA CCA CGC
gln his leu glu trp gly pro phe arg cys cys trp)OCH(pro gly gln pro ala pro arg
121/41   151/51   ⟵SEQ ID NO: 592
CTC GGT CGA GAC GTA TCG CGG CAG CGT TGG CCC TGT CGT TGC TGA CAA TTA CCG CTG GCC
leu gly arg asp val ser arg gln arg trp pro cys arg cys)OPA(gln leu pro leu ala
181/61   211/71   ⟵SEQ ID NO: 593
GCC GCA TAT TTG CCG CGC TGC CGC GGG CCG GAT C
ala ala tyr leu pro arg cys arg gly pro asp)

SEQ ID NOS:590-593

FIG. 34B

1/1   ⟋SEQ ID NO: 594   31/11
GAT CCA ACC TGC TGG GCC TGC GCC TTC GAA TCG ACG GCC AGG CCA CCG CTC GCT GCC GGC
(asp pro thr cys trp ala cys ala phe glu ser thr ala arg pro pro leu ala ala gly
61/21 ⟋SEQ ID NO: 595   91/31
AAC AAC ACC TGG AAT GGG GAC CTT TTC GGT GTT GCT GGT AAC CGG GAC AAC CGG CAC CAC
asn asn thr trp asn gly asp leu phe gly val ala gly asn arg asp asn arg his his
121/41   151/51
GCC TCG GTC GAG ACG TAT CGC GGC AGC GTT GGC CCT GTC GTT GCT GAC AAT TAC CGC TGG
ala ser val glu thr tyr arg gly ser val gly pro val val ala asp asn tyr arg trp
181/61   211/71
CCG CCG CAT ATT TGC CGC GCT GCC GCG GGC CGG ATC
pro pro his ile cys arg ala ala ala gly arg ile)

SEQ ID NOS:594-595

FIG. 34C

ORF according to Cole et al. (Nature 393:537-544) containing seq34A

1/1   SEQ ID NO: 596   31/11
```
tag ccg cag ggc cct gcg gct agg cgc ggc cgg tgc cgt tgg ccg cgg cgg caa tcg atg
AMB(pro gln gly pro ala ala arg arg gly arg cys arg trp pro arg arg gln ser met
```
61/21   SEQ ID NO: 597   91/31
```
ttg cag cag tta caa cgc caa atg gag tct gag cgc atc gtc gag ttc gat cag ctc ggc
leu gln gln leu gln arg gln met glu ser glu arg ile val glu phe asp gln leu gly
```
121/41   151/51
```
agg gga gac gtt gcg cag cga cgg atc caa cct gct ggg cct gcg cct tcg aat cga cgg
arg gly asp val ala gln arg arg ile gln pro ala gly pro ala pro ser asn arg arg
```
181/61   211/71
```
cca ggc cac cgc tcg ctg ccg gca aca aca cct gga atg ggg acc ttt tcg gtg ttg ctg
pro gly his arg ser leu pro ala thr thr pro gly met gly thr phe ser val leu leu
```
241/81   271/91
```
gta acc ggg aca acc ggc acc acg cct cgg tcg aga cgt atc gcg gca gcg ttg gcc ctg
val thr gly thr thr gly thr thr pro arg ser arg arg ile ala ala ala leu ala leu
```
301/101   331/111
```
tcg ttg ctg aca att acc gct ggc cgc cgc ata ttt gcc gcg ctg ccg cgg gcc gga tcc
ser leu leu thr ile thr ala gly arg arg ile phe ala ala leu pro arg ala gly ser
```
361/121   391/131
```
agg tcg acc tgc cag atc tca ccg cgc agc atc tac gcc gtt cgc tgc aaa ccg ccg act
arg ser thr cys gln ile ser pro arg ser ile tyr ala val arg cys lys pro pro thr
```
421/141   451/151
```
gcg acg gca ggc cca ctc tct tgg cat gcg tcc aat gct gcg acg tcc tcg gta gac aag
ala thr ala gly pro leu ser trp his ala ser asn ala ala thr ser ser val asp lys
```
481/161   511/171
```
ctc acg ctt ggc ttc atg ccg cag tcc tac cca tgt agt aac aga tag
leu thr leu gly phe met pro gln ser tyr pro cys ser asn arg)AMB
```

SEQ ID NOS:596-597

FIG. 34E

1/1   SEQ ID NO: 598   31/11
```
CAG TCT GTC GGC AAG GAG GGA CGC ATG CCA CTC TCC GAT CAT GAG CAG CGG ATG CTT GAC
(gln ser val gly lys glu gly arg met pro leu ser asp his glu gln arg met leu asp
```
61/21   SEQ ID NO: 599   91/31
```
CAG ATC GAG AGC GCT CTC TAC GCC GAA GAT CCC AAG TTC GCA TCG AGT GTC CGT GGC GGG
gln ile glu ser ala leu tyr ala glu asp pro lys phe ala ser ser val arg gly gly
```
121/41   151/51
```
GGC TTC CGC GCA CCG ACC GCG CGG CGG CGC CTG CAG GGC GCG GCG TTG TTC ATC ATC GGT
gly phe arg ala pro thr ala arg arg arg leu gln gly ala ala leu phe ile ile gly
```
181/61   211/71
```
CTG GGG ATG TTG GTT TCC GGC GTG GCG TTC AAA GAG ACC ATG ATC GGA AGT TTC CCG ATA
leu gly met leu val ser gly val ala phe lys glu thr met ile gly ser phe pro ile
```
241/81   271/91
```
CTC AGC GTT TTC GGT TTT GTC GTG ATG TTC GGT GGT GTG GTG TAT GCC ATC ACC GGT CCT
leu ser val phe gly phe val val met phe gly gly val val tyr ala ile thr gly pro
```
301/101   331/111
```
CGG TTG TCC GGC AGG ATG GAT CGT GGC GGA TCG GCT GCT GGG GCT TCG CGC CAG CGT CGT
arg leu ser gly arg met asp arg gly gly ser ala ala gly ala ser arg gln arg arg
```
361/121   391/131
```
ACC AAG GGG GCC GGG GGC TCA TTC ACC AGC CGT ATG GAA GAT C
thr lys gly ala gly gly ser phe thr ser arg met glu asp)
```

SEQ ID NOS:598-599

FIG. 35A

1/1 ⌐SEQ ID NO: 600   31/11
GAC AGT CTG TCG GCA AGG AGG GAC GCA TGC CAC TCT CCG ATC ATG AGC AGC GGA TGC TTG
(asp ser leu ser ala arg arg asp ala cys his ser pro ile met ser ser gly cys leu
61/21 ⌐ SEQ ID NO: 601   91/31
ACC AGA TCG AGA GCG CTC TCT ACG CCG AAG ATC CCA AGT TCG CAT CGA GTG TCC GTG GCG
thr arg ser arg ala leu ser thr pro lys ile pro ser ser his arg val ser val ala
121/41   151/51
GGG GCT TCC GCG CAC CGA CCG CGC GGC GGC GCC TGC AGG GCG CGG CGT TGT TCA TCA TCG
gly ala ser ala his arg pro arg gly gly ala cys arg ala arg arg cys ser ser ser
181/61   211/71
GTC TGG GGA TGT TGG TTT CCG GCG TGG CGT TCA AAG AGA CCA TGA TCG GAA GTT TCC CGA
val trp gly cys trp phe pro ala trp arg ser lys arg pro)OPA(ser glu val ser arg
241/81   271/91   ⌐SEQ ID NO: 602
TAC TCA GCG TTT TCG GTT TTG TCG TGA TGT TCG GTG GTG TGG TGT ATG CCA TCA CCG GTC
tyr ser ala phe ser val leu ser)OPA(cys ser val val trp cys met pro ser pro val
301/101   SEQ ID NO: 603⌐   331/111
CTC GGT TGT CCG GCA GGA TGG ATC GTG GCG GAT CGG CTG CTG GGG CTT CGC GCC AGC GTC
leu gly cys pro ala gly trp ile val ala asp arg leu leu gly leu arg ala ser val
361/121   391/131
GTA CCA AGG GGG CCG GGG GCT CAT TCA CCA GCC GTA TGG AAG ATC
val pro arg gly pro gly ala his ser pro ala val trp lys ile)

SEQ ID NOS:600-603

FIG. 35B

1/1 ⌐SEQ ID NO: 604   31/11
ACA GTC TGT CGG CAA GGA GGG ACG CAT GCC ACT CTC CGA TCA TGA GCA GCG GAT GCT TGA
(thr val cys arg gln gly gly thr his ala thr leu arg ser)OPA(ala ala asp ala)OPA
61/21 ⌐ SEQ ID NO: 605   91/31   SEQ ID NO: 606
CCA GAT CGA GAG CGC TCT CTA CGC CGA AGA TCC CAA GTT CGC ATC GAG TGT CCG TGG CGG
(pro asp arg glu arg ser leu arg arg arg ser gln val arg ile glu cys pro trp arg
121/41 ⌐ SEQ ID NO: 607   151/51
GGG CTT CCG CGC ACC GAC CGC GCG GCG GCG CCT GCA GGG CGC GGC GTT GTT CAT CAT CGG
gly leu pro arg thr asp arg ala ala ala pro ala gly arg gly val val his his arg
181/61   211/71
TCT GGG GAT GTT GGT TTC CGG CGT GGC GTT CAA AGA GAC CAT GAT CGG AAG TTT CCC GAT
ser gly asp val gly phe arg arg gly val gln arg asp his asp arg lys phe pro asp
241/81   271/91
ACT CAG CGT TTT CGG TTT TGT CGT GAT GTT CGG TGG TGT GGT GTA TGC CAT CAC CGG TCC
thr gln arg phe arg phe cys arg asp val arg trp cys gly val cys his his arg ser
301/101   331/111
TCG GTT GTC CGG CAG GAT GGA TCG TGG CGG ATC GGC TGC TGG GGC TTC GCG CCA GCG TCG
ser val val arg gln asp gly ser trp arg ile gly cys trp gly phe ala pro ala ser
361/121   391/131
TAC CAA GGG GGC CGG GGG CTC ATT CAC CAG CCG TAT GGA AGA TC
tyr gln gly gly arg gly leu ile his gln pro tyr gly arg)

SEQ ID NOS:604-607

FIG. 35C

sequence Rv2169c predicted by Cole et al. (Nature 393:537-544) and
partially containing seq35A

```
1/1  SEQ ID NO: 608                        31/11
atg cca ctc tcc gat cat gag cag cgg atg ctt gac cag atc gag agc gct ctc tac gcc
(Met pro leu ser asp his glu gln arg met leu asp gln ile glu ser ala leu tyr ala
61/21  SEQ ID NO: 609                      91/31
gaa gat ccc aag ttc gca tcg agt gtc cgt ggc ggg ggc ttc cgc gca ccg acc gcg cgg
glu asp pro lys phe ala ser ser val arg gly gly gly phe arg ala pro thr ala arg
121/41                                     151/51
cgg cgc ctg cag ggc gcg gcg ttg ttc atc atc ggt ctg ggg atg ttg gtt tcc ggc gtg
arg arg leu gln gly ala ala leu phe ile ile gly leu gly met leu val ser gly val
181/61                                     211/71
gcg ttc aaa gag acc atg atc gga agt ttc ccg ata ctc agc gtt ttc ggt ttt gtc gtg
ala phe lys glu thr met ile gly ser phe pro ile leu ser val phe gly phe val val
241/81                                     271/91
atg ttc ggt ggt gtg gtg tat gcc atc acc ggt cct cgg ttg tcc ggc agg atg gat cgt
met phe gly gly val val tyr ala ile thr gly pro arg leu ser gly arg met asp arg
301/101                                    331/111
ggc gga tcg gct gct ggg gct tcg cgc cag cgt cgt acc aag ggg gcc ggg ggc tca ttc
gly gly ser ala ala gly ala ser arg gln arg arg thr lys gly ala gly gly ser phe
361/121                                    391/131
acc agc cgt atg gaa gat cgg ttc cgg cgc cgc ttc gac gag taa
thr ser arg met glu asp arg phe arg arg arg phe asp glu)OCH
```

SEQ ID NOS:608-609

FIG. 35D

ORF according to Cole et al. (Nature 393:537-544) and containing Rv2169c

```
1/1  SEQ ID NO: 610                        31/11
tga cag tct gtc ggc aag gag gga cgc atg cca ctc tcc gat cat gag cag cgg atg ctt
OPA(gln ser val gly lys glu gly arg met pro leu ser asp his glu gln arg met leu
61/21  SEQ ID NO: 611                      91/31
gac cag atc gag agc gct ctc tac gcc gaa gat ccc aag ttc gca tcg agt gtc cgt ggc
asp gln ile glu ser ala leu tyr ala glu asp pro lys phe ala ser ser val arg gly
121/41                                     151/51
ggg ggc ttc cgc gca ccg acc gcg cgg cgc ctg cag ggc gcg gcg ttg ttc atc atc
gly gly phe arg ala pro thr ala arg arg arg leu gln gly ala ala leu phe ile ile
181/61                                     211/71
ggt ctg ggg atg ttg gtt tcc ggc gtg gcg ttc aaa gag acc atg atc gga agt ttc ccg
gly leu gly met leu val ser gly val ala phe lys glu thr met ile gly ser phe pro
241/81                                     271/91
ata ctc agc gtt ttc ggt ttt gtc gtg atg ttc ggt ggt gtg gtg tat gcc atc acc ggt
ile leu ser val phe gly phe val val met phe gly gly val val tyr ala ile thr gly
301/101                                    331/111
cct cgg ttg tcc ggc agg atg gat cgt ggc gga tcg gct gct ggg gct tcg cgc cag cgt
pro arg leu ser gly arg met asp arg gly gly ser ala ala gly ala ser arg gln arg
361/121                                    391/131
cgt acc aag ggg gcc ggg ggc tca ttc acc agc cgt atg gaa gat cgg ttc cgg cgc cgc
arg thr lys gly ala gly gly ser phe thr ser arg met glu asp arg phe arg arg arg
421/141
ttc gac gag taa
phe asp glu)OCH
```

SEQ ID NOS:610-611

FIG. 35E

1/1 ⟋SEQ ID NO: 612   31/11
GAC CTG GGA CGA AGA CGA CGG CAG CAG CCG CAA TCA GAT CTA CCC GGT CCT GGT CAA CGT
(asp leu gly arg arg arg arg gln gln pro gln ser asp leu pro gly pro gly gln arg
61/21 ⟋SEQ ID NO: 613   91/31
CAA TGG ACA CCC GAC TAC GGT GCG CCT GCG CGG CTC GAC AAT GCG CGG TTC CTG TTG CCC
gln trp thr pro asp tyr gly ala pro ala arg leu asp asn ala arg phe leu leu pro
121/41   151/51
GTG GTC GGA GTG CCA CCC GAC CAG GCC ACC GAC TTC GGC TCC GCT GTT GCA CCA GAA ACG
val val gly val pro pro asp gln ala thr asp phe gly ser ala val ala pro glu thr
181/61   211/71
ACG GCG CCG GTC TGG ATC ACC ATG CTG TGG CCG CTG GCC GAC CGG CCC CGG TTG GCC CCC
thr ala pro val trp ile thr met leu trp pro leu ala asp arg pro arg leu ala pro
241/81   271/91
GGG GCA CCC GGT GGC ACC GTT CCC GTC CGG CTG GTC GAC GAC GAC CTG GCA AAC TCG CTG
gly ala pro gly gly thr val pro val arg leu val asp asp asp leu ala asn ser leu
301/101   331/111
GCC AAC GGC GGC CGG CTG GAC ATC CTC CTG TCG GCG GCC GAG TTC GCC ACC AAC CGG GAA
ala asn gly gly arg leu asp ile leu leu ser ala ala glu phe ala thr asn arg glu
361/121   391/131
GTC GAC CCC GAC GGC GCC GTC GGC CGA GCG CTG TGC CTG GCC ATC GAC CCA GAT C
val asp pro asp gly ala val gly arg ala leu cys leu ala ile asp pro asp)

SEQ ID NOS:612-613

FIG. 36A

1/1 ⟋SEQ ID NO: 614   31/11
ACC TGG GAC GAA GAC GAC GGC AGC AGC CGC AAT CAG ATC TAC CCG GTC CTG GTC AAC GTC
(thr trp asp glu asp asp gly ser ser arg asn gln ile tyr pro val leu val asn val
61/21 ⟋SEQ ID NO: 615   91/31
AAT GGA CAC CCG ACT ACG GTG CGC CTG CGC GGC TCG ACA ATG CGC GGT TCC TGT TGC CCG
asn gly his pro thr thr val arg leu arg gly ser thr met arg gly ser cys cys pro
121/41   151/51
TGG TCG GAG TGC CAC CCG ACC AGG CCA CCG ACT TCG GCT CCG CTG TTG CAC CAG AAA CGA
trp ser glu cys his pro thr arg pro pro thr ser ala pro leu leu his gln lys arg
181/61   211/71
CGG CGC CGG TCT GGA TCA CCA TGC TGT GGC CGC TGG CCG ACC GGC CCC GGT TGG CCC CCG
arg arg arg ser gly ser pro cys cys gly arg trp pro thr gly pro gly trp pro pro
241/81   271/91
GGG CAC CCG GTG GCA CCG TTC CCG TCC GGC TGG TCG ACG ACG ACC TGG CAA ACT CGC TGG
gly his pro val ala pro phe pro ser gly trp ser thr thr thr trp gln thr arg trp
301/101   331/111
CCA ACG GCG GCC GGC TGG ACA TCC TCC TGT CGG CGG CCG AGT TCG CCA CCA ACC GGG AAG
pro thr ala ala gly trp thr ser ser cys arg arg pro ser ser pro pro thr gly lys
361/121   391/131
TCG ACC CCG ACG GCG CCG TCG GCC GAG CGC TGT GCC TGG CCA TCG ACC CAG ATC
ser thr pro thr ala pro ser ala glu arg cys ala trp pro ser thr gln ile)

SEQ ID NOS:614-615

FIG. 36B

1/1 ___SEQ ID NO: 616   31/11
CCT GGG ACG AAG ACG ACG GCA GCA GCC GCA ATC AGA TCT ACC CGG TCC TGG TCA ACG TCA
(pro gly thr lys thr thr ala ala ala ala ile arg ser thr arg ser trp ser thr ser
61/21 ___SEQ ID NO: 617   91/31
ATG GAC ACC CGA CTA CGG TGC GCC TGC GCG GCT CGA CAA TGC GCG GTT CCT GTT GCC CGT
met asp thr arg leu arg cys ala cys ala ala arg gln cys ala val pro val ala arg
121/41   151/51
GGT CGG AGT GCC ACC CGA CCA GGC CAC CGA CTT CGG CTC CGC TGT TGC ACC AGA AAC GAC
gly arg ser ala thr arg pro gly his arg leu arg leu arg cys cys thr arg asn asp
181/61   211/71
GGC GCC GGT CTG GAT CAC CAT GCT GTG GCC GCT GGC CGA CCG GCC CCG GTT GGC CCC CGG
gly ala gly leu asp his his ala val ala ala gly arg pro ala pro val gly pro arg
241/81   271/91
GGC ACC CGG TGG CAC CGT TCC CGT CCG GCT GGT CGA CGA CGA CCT GGC AAA CTC GCT GGC
gly thr arg trp his arg ser arg pro ala gly arg arg arg pro gly lys leu ala gly
301/101   331/111
CAA CGG CGG CCG GCT GGA CAT CCT CCT GTC GGC GGC CGA GTT CGC CAC CAA CCG GGA AGT
gln arg arg pro ala gly his pro pro val gly gly arg val arg his gln pro gly ser
361/121   391/131
CGA CCC CGA CGG CGC CGT CGG CCG AGC GCT GTG CCT GGC CAT CGA CCC AGA TC
arg pro arg arg arg arg arg pro ser ala val pro gly his arg pro arg)

SEQ ID NOS:616-617

FIG. 36C

Coding sequence Rv3909 predicted by Cole et al., 1998 (Nature 393
537-544) containing Seq 36A

1/1 ___SEQ ID NO: 618   31/11
GTG ACC GCA CTG CAA CTC GGC TGG GCC GCT TTG GCG CGC GTC ACC TCA GCG ATC GGC GTC
(met thr ala leu gln leu gly trp ala ala leu ala arg val thr ser ala ile gly val
61/21 ___SEQ ID NO: 619   91/31
GTG GCC GGC CTC GGG ATG GCG CTC ACG GTA CCG TCG GCG GCA CCG CAC GCG CTC GCA GGC
val ala gly leu gly met ala leu thr val pro ser ala ala pro his ala leu ala gly
121/41   151/51
GAG CCC AGC CCG ACG CCT TTT GTC CAG GTC CGC ATC GAT CAG GTG ACC CCG GAC GTG GTG
glu pro ser pro thr pro phe val gln val arg ile asp gln val thr pro asp val val
181/61   211/71
ACC ACT TCC AGC GAA CCC CAT GTC ACC GTC AGC GGA ACG GTG ACC AAT ACC GGT GAC CGC
thr thr ser ser glu pro his val thr val ser gly thr val thr asn thr gly asp arg
241/81   271/91
CCA GTC CGC GAT GTG ATG GTC CGG CTT GAG CAC GCC GCC GCG GTC ACG TCG TCA ACG GCG
pro val arg asp val met val arg leu glu his ala ala ala val thr ser ser thr ala
301/101   331/111
TTA CGC ACC TCG CTC GAC GGC GGC ACC GAC CAG TAC CAG CCG GCC GCG GAC TTC CTC ACG
leu arg thr ser leu asp gly gly thr asp gln tyr gln pro ala ala asp phe leu thr

SEQ ID NOS:618-619

FIG. 36D

361/121
GTC GCC CCC GAA CTA GAC CGC GGG CAA GAG
val ala pro glu leu asp arg gly gln glu
421/141
TCG CTG ACC AGG CCG TCG TTG GCC GTC AAC
ser leu thr arg pro ser leu ala val asn
481/161
GTC AAT GGG ACA CCC GAC TAC GGT GCG CCT
val asn gly thr pro asp tyr gly ala pro
541/181
CCC GTG GTC GGA GTG CCA CCC GAC CAG GCC
pro val val gly val pro pro asp gln ala
601/201
ACG ACG GCG CCG GTC TGG ATC ACC ATG CTG
thr thr ala pro val trp ile thr met leu
661/221
CCC GGG GCA CCC GGT GGC ACC GTT CCC GTC
pro gly ala pro gly gly thr val pro val
721/241
CTG GCC AAC GGC GGC CGG CTG GAC ATC CTC
leu ala asn gly gly arg leu asp ile leu
781/261
GAA GTC GAC CCC GAC GGC GCC GTC GGC CGA
glu val asp pro asp gly ala val gly arg
841/281
CTC ATC ACC GTC AAT GCG ATG ACC GGC GGC
leu ile thr val asn ala met thr gly gly
901/301
GCT CAA CTA CCG GGC ACC CCG ACC CAC CCG
ala gln leu pro gly thr pro thr his pro
961/321
CTG GAT CGA TTG CGG ACG CTA GTC CAC CGG
leu asp arg leu arg thr leu val his arg
1021/341
GCC GAC CTG GAT GCT TTG CAG CGG GTT AAT
ala asp leu asp ala leu gln arg val asn
1081/361
AGC CCC GCC GAC ATC GTC GAC CGC ATC CTG
ser pro ala asp ile val asp arg ile leu
1141/381
CTG CCC GAC GGC CCG TTG ACC GGC CGG GCG
leu pro asp gly pro leu thr gly arg ala
1201/401
GTT GCC GTC GCG GCC GCC GAT TTT AGC CCC
val ala val ala ala ala asp phe ser pro
1261/421
TCC GCG CTC TTA CCC GCT ACC GCG CCC CGG
ser ala leu leu pro ala thr ala pro arg
1321/441
TTT GAT CCC GCG GTC GGG GCC GCG CTG GCC
phe asp pro ala val gly ala ala leu ala
1381/461
TAT CTA GAT CCC TCG TTG TTC GTT CGG ATC
tyr leu asp pro ser leu phe val arg ile

391/131
GCC GGC TTT ACC CTC TCG GCC CCG CTG CGC
ala gly phe thr leu ser ala pro leu arg
451/151
CAG CCC GGG ATC TAC CCG GTC CTG GTC AAC
gln pro gly ile tyr pro val leu val asn
511/171
GCG CGG CTC GAC AAT GCG CGG TTC CTG TTG
ala arg leu asp asn ala arg phe leu leu
571/191
ACC GAC TTC GGC TCC GCT GTT GCA CCA GAA
thr asp phe gly ser ala val ala pro glu
631/211
TGG CCG CTG GCC GAC CGG CCC CGG TTG GCC
trp pro leu ala asp arg pro arg leu ala
691/231
CGG CTG GTC GAC GAC GAC CTG GCA AAC TCG
arg leu val asp asp asp leu ala asn ser
751/251
CTG TCG GCG GCC GAG TTC GCC ACC AAC CGG
leu ser ala ala glu phe ala thr asn arg
811/271
GCG CTG TGC CTG GCC ATC GAC CCA GAT CTA
ala leu cys leu ala ile asp pro asp leu
871/291
TAC GTC GTG TCC GAC TCG CCC GAC GGG GCC
tyr val val ser asp ser pro asp gly ala
931/311
GGC ACC GGC CAG GCC GCC GCA TCC AGC TGG
gly thr gly gln ala ala ala ser ser trp
991/331
ACA TGC GTG ACG CCG CTG CCT TTT GCC CAA
thr cys val thr pro leu pro phe ala gln
1051/351
GAT CCG AGG CTG AGC GCG ATC GCA ACC ATC
asp pro arg leu ser ala ile ala thr ile
1111/371
GAT GTC AGC TCC ACC CGC GGC GCA ACC GTG
asp val ser ser thr arg gly ala thr val
1171/391
ATC AAC TTG CTC AGC ACC CAC GGC AAC ACG
ile asn leu leu ser thr his gly asn thr
1231/411
GAG GAA CAG CAG GGT TCG TCC CAG ATC GGC
glu glu gln gln gly ser ser gln ile gly
1291/431
CGG TTG TCC CCG CGG GTG GTA GCG GCG CCG
arg leu ser pro arg val val ala ala pro
1351/451
GCC GCG GGA ACA AAC CCG ACC GTT CCT ACC
ala ala gly thr asn pro thr val pro thr
1411/471
GCG CAT GAA TCG ATC ACC GCG CGC CGC CAG
ala his glu ser ile thr ala arg arg gln

SEQ ID NOS:618-619(continued 1)

FIG. 36D (continued 1)

1441/481
GAC GCC TTG GGC GCA ATG CTG TGG CGC AGC
asp ala leu gly ala met leu trp arg ser
1501/501
ATC CTG GTG CCG CCG GCG TCG TGG AGC CTG
ile leu val pro pro ala ser trp ser leu
1561/521
GCG CTG GCC ACC GCC ATC CGG TCT GGC CTG
ala leu ala thr ala ile arg ser gly leu
1621/541
GCT GAC GCC GCG GCC CGC ACC GAG CCA CCG
ala asp ala ala ala arg thr glu pro pro
1681/561
GGC CGG TTC AAT GAC GAC ATC ACC ACG CAG
gly arg phe asn asp asp ile thr thr gln
1741/581
CTG ACC TCG GCG TTG ACC ATC GAT GAC CGC
leu thr ser ala leu thr ile asp asp arg
1801/601
CCA CTA CGC GAG GAC ATG TTG CGC GCG CTG
pro leu arg glu asp met leu arg ala leu
1861/621
GGG CTG GCC CAG CAG CGG CTG GCC GTC GTT
gly leu ala gln gln arg leu ala val val
1921/641
GTG ACC ATC GTC AAC CCG GGC GGC TCC TAC
val thr ile val asn pro gly gly ser tyr
1981/661
TTG GCG CTG CAT AAT GGC CTC GCC GTG CCA
leu ala leu his asn gly leu ala val pro
2041/681
CCC GGG ATG ACG GTG GCC GAT GTC GGT CAG
pro gly met thr val ala asp val gly gln
2101/701
CGA GTA CCA ATC GAG GTG AAC TTC ACA CAG
arg val pro ile glu val asn phe thr gln
2161/721
CCC GAC GGC GTC GCG CTG GGT GAA CCG GTG
pro asp gly val ala leu gly glu pro val
2221/741
AAG GTG TTG TTC GCG ATC ACG CTA TCC GCT
lys val leu phe ala ile thr leu ser ala
2281/761
CGC CTT TGG CAC CGG TTC CGT GGC CAG CCT
arg leu trp his arg phe arg gly gln pro
2341/781
CCT ACC GGC AAA CAC GCC CCG CAG CGC CGT
pro thr gly lys his ala pro gln arg arg
2401/801
CGG GTA TGA
arg val)OPA

1471/491
TTG GAG CCG AAT GCC GCG CCC CGT ACC CAA
leu glu pro asn ala ala pro arg thr gln
1531/511
GCC AGC GAC GAC GCG CAG GTC ATC CTG ACC
ala ser asp asp ala gln val ile leu thr
1591/531
GCC GTG CCG CGA CCA CTA CCG GCG GTG ATC
ala val pro arg pro leu pro ala val ile
1651/551
GAA CCC CCG GGC GCT TAC AGC GCC GCT CGC
glu pro pro gly ala tyr ser ala ala arg
1711/571
ATC GGC GGG CAG GTT GCC CGG CTA TGG AAG
ile gly gly gln val ala arg leu trp lys
1771/591
ACC GGG CTG ACC GGC GTG CAG TAC ACC GCA
thr gly leu thr gly val gln tyr thr ala
1831/611
AGC CAA TCG CTA CCA CCC GAT ACC CGC AAC
ser gln ser leu pro pro asp thr arg asn
1891/631
GGA AAG ACG ATC GAC GAT CTT TTC GGC GCG
gly lys thr ile asp asp leu phe gly ala
1951/651
ACT CTG GCC ACC GAG CAC AGT CCG CTG CCG
thr leu ala thr glu his ser pro leu pro
2011/671
ATC CGG GTC CGG CTA CAG GTC GAT GCT CCG
ile arg val arg leu gln val asp ala pro
2071/691
ATC GAG CTA CCG CCC GGG TAC CTG CCG CTA
ile glu leu pro pro gly tyr leu pro leu
2131/711
CGG GTT GCC GTC GAC GTG TCG CTG CGG ACC
arg val ala val asp val ser leu arg thr
2191/731
CGG TTG TCG GTG CAC TCC AAC GCC TAC GGC
arg leu ser val his ser asn ala tyr gly
2251/751
GCG GCC GTG CTG GTA ACG CTG GCG GGC CGG
ala ala val leu val thr leu ala gly arg
2311/771
GAT CGC GCC GAC CTG GAT CGC CCC GAC CTG
asp arg ala asp leu asp arg pro asp leu
2371/791
GCC GTA GCC AGT CGG GAT GAC GAA AAG CAC
ala val ala ser arg asp asp glu lys his

SEQ ID NOS:618-619 (continued 2)

FIG. 36D (continued 2)

ORF according to Cole et al., 1998 (Nature 393 537-544) and containing Rv 3909.

1/1   SEQ ID NO: 620   31/11
TGA CTC AGC ACC GGG TCA GCA CAA CGG TCC CGG GCC GGG GCC GTG ACC GCA CTG CAA CTC
OPA(leu ser thr gly ser ala gln arg ser arg ala gly ala val thr ala leu gln leu
61/21   SEQ ID NO: 621   91/31
GGC TGG GCC GCT TTG GCG CGC GTC ACC TCA GCG ATC GGC GTC GTG GCC GGC CTC GGG ATG
gly trp ala ala leu ala arg val thr ser ala ile gly val val ala gly leu gly met
121/41   151/51
GCG CTC ACG GTA CCG TCG GCG GCA CCG CAC GCG CTC GCA GGC GAG CCC AGC CCG ACG CCT
ala leu thr val pro ser ala ala pro his ala leu ala gly glu pro ser pro thr pro
181/61   211/71
TTT GTC CAG GTC CGC ATC GAT CAG GTG ACC CCG GAC GTG GTG ACC ACT TCC AGC GAA CCC
phe val gln val arg ile asp gln val thr pro asp val val thr thr ser ser glu pro
241/81   271/91
CAT GTC ACC GTC AGC GGA ACG GTG ACC AAT ACC GGT GAC CGC CCA GTC CGC GAT GTG ATG
his val thr val ser gly thr val thr asn thr gly asp arg pro val arg asp val met
301/101   331/111
GTC CGG CTT GAG CAC GCC GCC GCG GTC ACG TCG TCA ACG GCG TTA CGC ACC TCG CTC GAC
val arg leu glu his ala ala ala val thr ser ser thr ala leu arg thr ser leu asp
361/121   391/131
GGC GGC ACC GAC CAG TAC CAG CCG GCC GCG GAC TTC CTC ACG GTC GCC CCC GAA CTA GAC
gly gly thr asp gln tyr gln pro ala ala asp phe leu thr val ala pro glu leu asp
421/141   451/151
CGC GGG CAA GAG GCC GGC TTT ACC CTC TCG GCC CCG CTG CGC TCG CTG ACC AGG CCG TCG
arg gly gln glu ala gly phe thr leu ser ala pro leu arg ser leu thr arg pro ser
481/161   511/171
TTG GCC GTC AAC CAG CCC GGG ATC TAC CCG GTC CTG GTC AAC GTC AAT GGG ACA CCC GAC
leu ala val asn gln pro gly ile tyr pro val leu val asn val asn gly thr pro asp
541/181   571/191
TAC GGT GCG CCT GCG CGG CTC GAC AAT GCG CGG TTC CTG TTG CCC GTG GTC GGA GTG CCA
tyr gly ala pro ala arg leu asp asn ala arg phe leu leu pro val val gly val pro
601/201   631/211
CCC GAC CAG GCC ACC GAC TTC GGC TCC GCT GTT GCA CCA GAA ACG ACG GCG CCG GTC TGG
pro asp gln ala thr asp phe gly ser ala val ala pro glu thr thr ala pro val trp
661/221   691/231
ATC ACC ATG CTG TGG CCG CTG GCC GAC CGG CCC CGG TTG GCC CCC GGG GCA CCC GGT GGC
ile thr met leu trp pro leu ala asp arg pro arg leu ala pro gly ala pro gly gly
721/241   751/251
ACC GTT CCC GTC CGG CTG GTC GAC GAC GAC CTG GCA AAC TCG CTG GCC AAC GGC GGC CGG
thr val pro val arg leu val asp asp asp leu ala asn ser leu ala asn gly gly arg
781/261   811/271
CTG GAC ATC CTC CTG TCG GCG GCC GAG TTC GCC ACC AAC CGG GAA GTC GAC CCC GAC GGC
leu asp ile leu leu ser ala ala glu phe ala thr asn arg glu val asp pro asp gly
841/281   871/291
GCC GTC GGC CGA GCG CTG TGC CTG GCC ATC GAC CCA GAT CTA CTC ATC ACC GTC AAT GCG
ala val gly arg ala leu cys leu ala ile asp pro asp leu leu ile thr val asn ala
901/301   931/311
ATG ACC GGC GGC TAC GTC GTG TCC GAC TCG CCC GAC GGG GCC GCT CAA CTA CCG GGC ACC
met thr gly gly tyr val val ser asp ser pro asp gly ala ala gln leu pro gly thr
961/321   991/331
CCG ACC CAC CCG GGC ACC GGC CAG GCC GCC GCA TCC AGC TGG CTG GAT CGA TTG CGG ACG
pro thr his pro gly thr gly gln ala ala ala ser ser trp leu asp arg leu arg thr

SEQ ID NOS:620-621

FIG. 36E

1021/341
CTA GTC CAC CGG ACA TGC GTG ACG CCG CTG
leu val his arg thr cys val thr pro leu

1051/351
CCT TTT GCC CAA GCC GAC CTG GAT GCT TTG
pro phe ala gln ala asp leu asp ala leu

1081/361
CAG CGG GTT AAT GAT CCG AGG CTG AGC GCG
gln arg val asn asp pro arg leu ser ala

1111/371
ATC GCA ACC ATC AGC CCC GCC GAC ATC GTC
ile ala thr ile ser pro ala asp ile val

1141/381
GAC CGC ATC CTG GAT GTC AGC TCC ACC CGC
asp arg ile leu asp val ser ser thr arg

1171/391
GGC GCA ACC GTG CTG CCC GAC GGC CCG TTG
gly ala thr val leu pro asp gly pro leu

1201/401
ACC GGC CGG GCG ATC AAC TTG CTC AGC ACC
thr gly arg ala ile asn leu leu ser thr

1231/411
CAC GGC AAC ACG GTT GCC GTC GCG GCC GCC
his gly asn thr val ala val ala ala ala

1261/421
GAT TTT AGC CCC GAG GAA CAG CAG GGT TCG
asp phe ser pro glu glu gln gln gly ser

1291/431
TCC CAG ATC GGC TCC GCG CTC TTA CCC GCT
ser gln ile gly ser ala leu leu pro ala

1321/441
ACC GCG CCC CGG CGG TTG TCC CCG CGG GTG
thr ala pro arg arg leu ser pro arg val

1351/451
GTA GCG GCG CCG TTT GAT CCC GCG GTC GGG
val ala ala pro phe asp pro ala val gly

1381/461
GCC GCG CTG GCC GCC GCG GGA ACA AAC CCG
ala ala leu ala ala ala gly thr asn pro

1411/471
ACC GTT CCT ACC TAT CTA GAT CCC TCG TTG
thr val pro thr tyr leu asp pro ser leu

1441/481
TTC GTT CGG ATC GCG CAT GAA TCG ATC ACC
phe val arg ile ala his glu ser ile thr

1471/491
GCG CGC CGC CAG GAC GCC TTG GGC GCA ATG.
ala arg arg gln asp ala leu gly ala met

1501/501
CTG TGG CGC AGC TTG GAG CCG AAT GCC GCG
leu trp arg ser leu glu pro asn ala ala

1531/511
CCC CGT ACC CAA ATC CTG GTG CCG CCG GCG
pro arg thr gln ile leu val pro pro ala

1561/521
TCG TGG AGC CTG GCC AGC GAC GAC GCG CAG
ser trp ser leu ala ser asp asp ala gln

1591/531
GTC ATC CTG ACC GCG CTG GCC ACC GCC ATC
val ile leu thr ala leu ala thr ala ile

1621/541
CGG TCT GGC CTG GCC GTG CCG CGA CCA CTA
arg ser gly leu ala val pro arg pro leu

1651/551
CCG GCG GTG ATC GCT GAC GCC GCG GCC CGC
pro ala val ile ala asp ala ala ala arg

1681/561
ACC GAG CCA CCG GAA CCC CCG GGC GCT TAC
thr glu pro pro glu pro pro gly ala tyr

1711/571
AGC GCC GCT CGC GGC CGG TTC AAT GAC GAC
ser ala ala arg gly arg phe asn asp asp

1741/581
ATC ACC ACG CAG ATC GGC GGG CAG GTT GCC
ile thr thr gln ile gly gly gln val ala

1771/591
CGG CTA TGG AAG CTG ACC TCG GCG TTG ACC
arg leu trp lys leu thr ser ala leu thr

1801/601
ATC GAT GAC CGC ACC GGG CTG ACC GGC GTG
ile asp asp arg thr gly leu thr gly val

1831/611
CAG TAC ACC GCA CCA CTA CGC GAG GAC ATG
gln tyr thr ala pro leu arg glu asp met

1861/621
TTG CGC GCG CTG AGC CAA TCG CTA CCA CCC
leu arg ala leu ser gln ser leu pro pro

1891/631
GAT ACC CGC AAC GGG CTG GCC CAG CAG CGG
asp thr arg asn gly leu ala gln gln arg

1921/641
CTG GCC GTC GTT GGA AAG ACG ATC GAC GAT
leu ala val val gly lys thr ile asp asp

1951/651
CTT TTC GGC GCG GTG ACC ATC GTC AAC CCG
leu phe gly ala val thr ile val asn pro

1981/661
GGC GGC TCC TAC ACT CTG GCC ACC GAG CAC
gly gly ser tyr thr leu ala thr glu his

2011/671
AGT CCG CTG CCG TTG GCG CTG CAT AAT GGC
ser pro leu pro leu ala leu his asn gly

SEQ ID NOS: 620-621 (continued 1)

FIG. 36E (continued 1)

2041/681   2071/691
CTC GCC GTG CCA ATC CGG GTC CGG CTA CAG GTC GAT GCT CCG CCC GGG ATG ACG GTG GCC
leu ala val pro ile arg val arg leu gln val asp ala pro pro gly met thr val ala
2101/701   2131/711
GAT GTC GGT CAG ATC GAG CTA CCG CCC GGG TAC CTG CCG CTA CGA GTA CCA ATC GAG GTG
asp val gly gln ile glu leu pro pro gly tyr leu pro leu arg val pro ile glu val
2161/721   2191/731
AAC TTC ACA CAG CGG GTT GCC GTC GAC GTG TCG CTG CGG ACC CCC GAC GGC GTC GCG CTG
asn phe thr gln arg val ala val asp val ser leu arg thr pro asp gly val ala leu
2221/741   2251/751
GGT GAA CCG GTG CGG TTG TCG GTG CAC TCC AAC GCC TAC GGC AAG GTG TTG TTC GCG ATC
gly glu pro val arg leu ser val his ser asn ala tyr gly lys val leu phe ala ile
2281/761   2311/771
ACG CTA TCC GCT GCG GCC GTG CTG GTA ACG CTG GCG GGC CGG CGC CTT TGG CAC CGG TTC
thr leu ser ala ala ala val leu val thr leu ala gly arg arg leu trp his arg phe
2341/781   2371/791
CGT GGC CAG CCT GAT CGC GCC GAC CTG GAT CGC CCC GAC CTG CCT ACC GGC AAA CAC GCC
arg gly gln pro asp arg ala asp leu asp arg pro asp leu pro thr gly lys his ala
2401/801   2431/811
CCG CAG CGC CGT GCC GTA GCC AGT CGG GAT GAC GAA AAG CAC CGG GTA TGA
pro gln arg arg ala val ala ser arg asp asp glu lys his arg val)OPA

SEQ ID NOS: 620-621 (continued 2)

FIG. 36E (continued 2)

1/1 ⟋SEQ ID NO: 622   31/11
ATC CGC GCG TTG GCG TCG CAT CCG AAC ATC GTC GGA GTC AAG GAC GCC AAA GCC GAC CTG
(ile arg ala leu ala ser his pro asn ile val gly val lys asp ala lys ala asp leu
61/21 ⟍SEQ ID NO: 623   91/31
CAC AGC GGC GCC CAA ATC ATG GCC GAC ACC GGA CTG GCC TAC TAT TCC GGC GAC GAC GCG
his ser gly ala gln ile met ala asp thr gly leu ala tyr tyr ser gly asp asp ala
121/41   151/51
CTC AAC CTG CCC TGG CTG GCC ATG GGC GCC ACG GGC TTC ATC AGC GTG ATT GCC CAC CTG
leu asn leu pro trp leu ala met gly ala thr gly phe ile ser val ile ala his leu
181/61   211/71
GCA GCC GGG CAG CTT CGA GAG TTG TTG TCC GCC TTC GGT TCT GGG GAT ATC GCC ACC GCC
ala ala gly gln leu arg glu leu leu ser ala phe gly ser gly asp ile ala thr ala
241/81
CGC AAG ATC
arg lys ile)

SEQ ID NOS:622-623

FIG. 37A

1/1 ——SEQ ID NO: 624   31/11
GAT CCG CGC GTT GGC GTC GCA TCC GAA CAT CGT CGG AGT CAA GGA CGC CAA AGC CGA CCT
(asp pro arg val gly val ala ser glu his arg arg ser gln gly arg gln ser arg pro
61/21 ——SEQ ID NO: 625   91/31
GCA CAG CGG CGC CCA AAT CAT GGC CGA CAC CGG ACT GGC CTA CTA TTC CGG CGA CGA CGC
ala gln arg arg pro asn his gly arg his arg thr gly leu leu phe arg arg arg arg
121/41   151/51
GCT CAA CCT GCC CTG GCT GGC CAT GGG CGC CAC GGG CTT CAT CAG CGT GAT TGC CCA CCT
ala gln pro ala leu ala gly his gly arg his gly leu his gln arg asp cys pro pro
181/61   211/71
GGC AGC CGG GCA GCT TCG AGA GTT GTT GTC CGC CTT CGG TTC TGG GGA TAT CGC CAC CGC
gly ser arg ala ala ser arg val val val arg leu arg phe trp gly tyr arg his arg
241/81
CCG CAA GAT C
pro gln asp)

SEQ ID NOS:624-625

FIG. 37B

1/1 ——SEQ ID NO: 626   31/11
TCC GCG CGT TGG CGT CGC ATC CGA ACA TCG TCG GAG TCA AGG ACG CCA AAG CCG ACC TGC
(ser ala arg trp arg arg ile arg thr ser ser glu ser arg thr pro lys pro thr cys
61/21 ——SEQ ID NO: 627   91/31
ACA GCG GCG CCC AAA TCA TGG CCG ACA CCG GAC TGG CCT ACT ATT CCG GCG ACG ACG CGC
thr ala ala pro lys ser trp pro thr pro asp trp pro thr ile pro ala thr thr arg
121/41   151/51
TCA ACC TGC CCT GGC TGG CCA TGG GCG CCA CGG GCT TCA TCA GCG TGA TTG CCC ACC TGG
ser thr cys pro gly trp pro trp ala pro arg ala ser ser ala)OPA(leu pro thr trp
181/61   211/71 SEQ ID NO: 628
CAG CCG GGC AGC TTC GAG AGT TGT TGT CCG CCT TCG GTT CTG GGG ATA TCG CCA CCG CCC
gln pro gly ser phe glu ser cys cys pro pro ser val leu gly ile ser pro pro pro
241/81
GCA AGA TC
ala arg)

SEQ ID NOS:626-628

FIG. 37C

Coding sequence Rv2753c predicted by Cole et al., 1998 (Nature 393 537-544) containing Seq 37A

1/1 ──SEQ ID NO: 629   31/11
GTG ACC ACC GTC GGA TTC GAC GTC GCA GCG CGC CTA GGA ACC CTG CTG ACC GCG ATG GTG
(val thr thr val gly phe asp val ala ala arg leu gly thr leu leu thr ala met val
61/21 ──SEQ ID NO: 630   91/31
ACA CCG TTT AGC GGC GAT GGC TCC CTG GAC ACC GCC ACC GCG GCG CGG CTG GCC AAC CAC
thr pro phe ser gly asp gly ser leu asp thr ala thr ala ala arg leu ala asn his
121/41   151/51
CTG GTC GAT CAG GGG TGC GAC GGT CTG GTG GTC TCG GGC ACC ACC GGC GAG TCG CCG ACC
leu val asp gln gly cys asp gly leu val val ser gly thr thr gly glu ser pro thr
181/61   211/71
ACC ACC GAC GGG GAG AAA ATC GAG CTG CTG CGG GCC GTC TTG GAA GCG GTG GGG GAC CGG
thr thr asp gly glu lys ile glu leu leu arg ala val leu glu ala val gly asp arg
241/81   271/91
GCC CGT GTT ATC GCC GGT GCC GGC ACC TAT GAC ACC GCG CAC AGC ATC CGG CTG GCC AAG
ala arg val ile ala gly ala gly thr tyr asp thr ala his ser ile arg leu ala lys
301/101   331/111
GCT TGT GCG GCC GAG GGT GCG CAC GGG CTG CTG GTG GTC ACG CCC TAC TAT TCC AAG CCG
ala cys ala ala glu gly ala his gly leu leu val val thr pro tyr tyr ser lys pro
361/121   391/131
CCG CAG CGG GGG CTG CAA GCC CAT TTC ACC GCC GTC GCC GAC GCG ACC GAG CTG CCG ATG
pro gln arg gly leu gln ala his phe thr ala val ala asp ala thr glu leu pro met
421/141   451/151
CTG CTC TAT GAC ATC CCG GGG CGG TCG GCG GTG CCG ATC GAG CCC GAC ACG ATC CGC GCG
leu leu tyr asp ile pro gly arg ser ala val pro ile glu pro asp thr ile arg ala
481/161   511/171
TTG GCG TCG CAT CCG AAC ATC GTC GGA GTC AAG GAC GCC AAA GCC GAC CTG CAC AGC GGC
leu ala ser his pro asn ile val gly val lys asp ala lys ala asp leu his ser gly
541/181   571/191
GCC CAA ATC ATG GCC GAC ACC GGA CTG GCC TAC TAT TCC GGC GAC GAC GCG CTC AAC CTG
ala gln ile met ala asp thr gly leu ala tyr tyr ser gly asp asp ala leu asn leu
601/201   631/211
CCC TGG CTG GCC ATG GGC GCC ACG GGC TTC ATC AGC GTG ATT GCC CAC CTG GCA GCC GGG
pro trp leu ala met gly ala thr gly phe ile ser val ile ala his leu ala ala gly
661/221   691/231
CAG CTT CGA GAG TTG TTG TCC GCC TTC GGT TCT GGG GAT ATC GCC ACC GCC CGC AAG ATC
gln leu arg glu leu leu ser ala phe gly ser gly asp ile ala thr ala arg lys ile
721/241   751/251
AAC ATT GCG GTC GCC CCG CTG TGC AAC GCG ATG AGC CGC CTG GGT GGG GTG ACG TTG TCC
asn ile ala val ala pro leu cys asn ala met ser arg leu gly gly val thr leu ser
781/261   811/271
AAG GCG GGC TTG CGG CTG CAG GGC ATC GAC GTC GGT GAT CCC CGG CTG CCC CAG GTG GCC
lys ala gly leu arg leu gln gly ile asp val gly asp pro arg leu pro gln val ala
841/281   871/291
GCG ACA CCG GAG CAG ATC GAC GCG TTG GCC GCC GAC ATG CGC GCG GCC TCG GTG CTT CGG
ala thr pro glu gln ile asp ala leu ala ala asp met arg ala ala ser val leu arg)

901/301
TGA
OPA

SEQ ID NOS:629-630

FIG. 37D

ORF according to Cole et al., 1998 (Nature 393 537-544) containing Rv2753c

1/1 — SEQ ID NO: 631   31/11
TAA GGT GAG CGC CGT GGC CGA GAC CGC GCC GCT GCG CGT GCA ACT GAT CGC CAA GAC CGA
OCH(gly glu arg arg gly arg asp arg ala ala ala arg ala thr asp arg gln asp arg
61/21 — SEQ ID NO: 632   91/31
CTT CTT GGC CCC ACC CGA CGT GCC CTG GAC CAC CGA CGC CGA CGG CGG ACC CGC GCT GGT
leu leu gly pro thr arg arg ala leu asp his arg arg arg arg thr arg ala gly
121/41   151/51
CGA GTT CGC CGG CCG GGC CTG CTA TCA GAG CTG GTC CAA GCC CAA TCC CAA GAC CGC CAC
arg val arg arg pro gly leu leu ser glu leu val gln ala gln ser gln asp arg his
181/61   211/71
CAA CGC CGG CTA CCT CCG GCA CAT CAT CGA CGT CGG ACA TTT CTC GGT GCT AGA GCA TGC
gln arg arg leu pro pro ala his his arg arg arg thr phe leu gly ala arg ala cys
241/81   271/91
CAG CGT GTC GTT CTA CAT CAC CGG GAT CTC GCG ATC GTG CAC CCA CGA GCT GAT CCG CCA
gln arg val val leu his his arg asp leu ala ile val his pro arg ala asp pro pro
301/101   331/111
CCG GCA TTT CTC CTA CTC GCA GCT CTC CCA GCG CTA CGT ACC CGA GAA GGA CTC GCG GGT
pro ala phe leu leu leu ala ala leu pro ala leu arg thr arg glu gly leu ala gly
361/121   391/131
CGT CGT GCC GCC CGG CAT GGA GGA CGA CGC CGA CCT GCG CCA CAT CCT GAC CGA GGC CGC
arg arg ala ala arg his gly gly arg arg arg pro ala pro his pro asp arg gly arg
421/141   451/151
CGA CGC CGC CCG CGC CAC CTA CAG CGA GCT GCT GGC CAA GCT GGA AGC CAA GTT CGC CGA
arg arg arg pro arg his leu gln arg ala ala gly gln ala gly ser gln val arg arg
481/161   511/171
CCA ACC CAA CGC GAT CCT GCG CCG CAA GCA GGC CCG CCA AGC CGC CCG CGC GGT GCT GCC
pro thr gln arg asp pro ala pro gln ala gly pro pro ser arg pro arg gly ala ala
541/181   571/191
CAA CGC CAC CGA AAC CCG CAT CGT GGT GAC CGG CAA CTA CCG GGC CTG GCG GCA CTT CAT
gln arg his arg asn pro his arg gly asp arg gln leu pro gly leu ala ala leu his
601/201   631/211
CGC AAT GCG GGC CAG CGA GCA CGC CGA CGT GGA AAT CCG GCG ACT GGC CAT CGA ATG CCT
arg asn ala gly gln arg ala arg arg arg gly asn pro ala thr gly his arg met pro
661/221   691/231
GCG CCA GCT CGC CGC CGT GGC CCC CGC GGT GTT CGC CGA CTT CGA GGT GAC CAC CCT GGC
ala pro ala arg arg arg gly pro arg gly val arg arg leu arg gly asp his pro gly
721/241   751/251
CGA CGG CAC CGA GGT GGC GAC CAG CCC GTT GGC GAC CGA AGC CTG AGG CGG CGT GTC GCT
arg arg his arg gly gly asp gln pro val gly asp arg ser leu arg arg arg val ala
781/261   811/271
GGA CAA ACA CGC GCG CTC GCG GCC GGG ATA AAG CGC CAG GTA ACC TTG GGA GCC GTG ACC
gly gln thr arg ala leu ala ala gly ile lys arg gln val thr leu gly ala val thr
841/281   871/291
ACC GTC GGA TTC GAC GTC GCA GCG CGC CTA GGA ACC CTG CTG ACC GCG ATG GTG ACA CCG
thr val gly phe asp val ala ala arg leu gly thr leu leu thr ala met val thr pro
901/301   931/311
TTT AGC GGC GAT GGC TCC CTG GAC ACC GCC ACC GCG GCG CGG CTG GCC AAC CAC CTG GTC
phe ser gly asp gly ser leu asp thr ala thr ala ala arg leu ala asn his leu val
961/321   991/331
GAT CAG GGG TGC GAC GGT CTG GTG GTC TCG GGC ACC ACC GGC GAG TCG CCG ACC ACC ACC
asp gln gly cys asp gly leu val val ser gly thr thr gly glu ser pro thr thr thr

SEQ ID NOS:631-632

FIG. 37E

1021/341
GAC GGG GAG AAA ATC GAG CTG CTG CGG GCC
asp gly glu lys ile glu leu leu arg ala
1081/361
GTT ATC GCC GGT GCC GGC ACC TAT GAC ACC
val ile ala gly ala gly thr tyr asp thr
1141/381
GCG GCC GAG GGT GCG CAC GGG CTG CTG GTG
ala ala glu gly ala his gly leu leu val
1201/401
CGG GGG CTG CAA GCC CAT TTC ACC GCC GTC
arg gly leu gln ala his phe thr ala val
1261/421
TAT GAC ATC CCG GGG CGG TCG GCG GTG CCG
tyr asp ile pro gly arg ser ala val pro
1321/441
TCG CAT CCG AAC ATC GTC GGA GTC AAG GAC
ser his pro asn ile val gly val lys asp
1381/461
ATC ATG GCC GAC ACC GGA CTG GCC TAC TAT
ile met ala asp thr gly leu ala tyr tyr
1441/481
CTG GCC ATG GGC GCC ACG GGC TTC ATC AGC
leu ala met gly ala thr gly phe ile ser
1501/501
CGA GAG TTG TTG TCC GCC TTC GGT TCT GGG
arg glu leu leu ser ala phe gly ser gly
1561/521
GCG GTC GCC CCG CTG TGC AAC GCG ATG AGC
ala val ala pro leu cys asn ala met ser
1621/541
GGC TTG CGG CTG CAG GGC ATC GAC GTC GGT
gly leu arg leu gln gly ile asp val gly
1681/561
CCG GAG CAG ATC GAC GCG TTG GCC GCC GAC
pro glu gln ile asp ala leu ala ala asp

1051/351
GTC TTG GAA GCG GTG GGG GAC CGG GCC CGT
val leu glu ala val gly asp arg ala arg
1111/371
GCG CAC AGC ATC CGG CTG GCC AAG GCT TGT
ala his ser ile arg leu ala lys ala cys
1171/391
GTC ACG CCC TAC TAT TCC AAG CCG CCG CAG
val thr pro tyr tyr ser lys pro pro gln
1231/411
GCC GAC GCG ACC GAG CTG CCG ATG CTG CTC
ala asp ala thr glu leu pro met leu leu
1291/431
ATC GAG CCC GAC ACG ATC CGC GCG TTG GCG
ile glu pro asp thr ile arg ala leu ala
1351/451
GCC AAA GCC GAC CTG CAC AGC GGC GCC CAA
ala lys ala asp leu his ser gly ala gln
1411/471
TCC GGC GAC GAC GCG CTC AAC CTG CCC TGG
ser gly asp asp ala leu asn leu pro trp
1471/491
GTG ATT GCC CAC CTG GCA GCC GGG CAG CTT
val ile ala his leu ala ala gly gln leu
1531/511
GAT ATC GCC ACC GCC CGC AAG ATC AAC ATT
asp ile ala thr ala arg lys ile asn ile
1591/531
CGC CTG GGT GGG GTG ACG TTG TCC AAG GCG
arg leu gly gly val thr leu ser lys ala
1651/551
GAT CCC CGG CTG CCC CAG GTG GCC GCG ACA
asp pro arg leu pro gln val ala ala thr
1711/571
ATG CGC GCG GCC TCG GTG CTT CGG TGA
met arg ala ala ser val leu arg)OPA

SEQ ID NOS:631-632 (continued 1)

FIG. 37E (continued 1)

1/1 ────SEQ ID NO: 633   31/11
GCG GTG AAC TGG TGG GCC CGG ATG GTT CAA   GTA CGC CGT CGC AAA CTC GAG CAC AAC AGG
(ala val asn trp trp ala arg met val gln   val arg arg arg lys leu glu his asn arg

61/21   ────SEQ ID NO: 634   91/31
AGA CGA CGG ATG GAA GGA GAT GCT GGC GCC   GGC CAG CTG AAC CCT GCC GAT GCG AAT AAG
arg arg arg met glu gly asp ala gly ala   gly gln leu asn pro ala asp ala asn lys

121/41   151/51
TCG TCG TCT ACG GAG GTG AAG GCG GCG GAT   TCG GCG GAA TCT GAC GCC GGA GCC GAC CAG
ser ser ser thr glu val lys ala ala asp   ser ala glu ser asp ala gly ala asp gln

181/61   211/71
ACT GGC CCG CAG GTG AAG GCG GCG GAT TCG   GCG GAA TCT GAC GCC GGA GAG CTC GGC GAG
thr gly pro gln val lys ala ala asp ser   ala glu ser asp ala gly glu leu gly glu

241/81   271/91
GAC GCG TGC CCA GAA CAG GCC CTC GTC GAG   CGG CGC CCG TCG CGG TTG CGG CGA GGC TGG
asp ala cys pro glu gln ala leu val glu   arg arg pro ser arg leu arg arg gly trp

301/101   331/111
CTT GTT GGC ATT GCG GCG ACG CTG CTC GCG   TTG GCC GGT GGC CTT GGC GCA GCG GGT TAT
leu val gly ile ala ala thr leu leu ala   leu ala gly gly leu gly ala ala gly tyr

361/121   391/131
TTT GCG TTG CGC TCA CAC CAG GAA AGC CAA   TCA ATC GCG CGC GAG GAC CTT GCG GCC ATT
phe ala leu arg ser his gln glu ser gln   ser ile ala arg glu asp leu ala ala ile

421/141   451/151
GAG GCC GCT AAG GAT TGC GTT GCG GCC ACG   CAG GCA CCC GAT GCT GGG GCG ATG TCG GCT
glu ala ala lys asp cys val ala ala thr   gln ala pro asp ala gly ala met ser ala

481/161
AGC ATG CAG AAG ATC
ser met gln lys ile)

SEQ ID NOS:633-634

FIG. 38A

1/1 ────SEQ ID NO: 635   31/11
CAG CGG TGA ACT GGT GGG CCC GGA TGG TTC   AAG TAC GCC GTC GCA AAC TCG AGC ACA ACA
gln arg OPA(thr gly gly pro gly trp phe   lys tyr ala val ala asn ser ser thr thr

61/21   └──SEQ ID NO: 636   91/31
GGA GAC GAC GGA TGG AAG GAG ATG CTG GCG   CCG GCC AGC TGA ACC CTG CCG ATG CGA ATA
gly asp asp gly trp lys glu met leu ala   pro ala ser)OPA(thr leu pro met arg ile

121/41   151/51   └──SEQ ID NO: 637
AGT CGT CGT CTA CGG AGG TGA AGG CGG CGG   ATT CGG CGG AAT CTG ACG CCG GAG CCG ACC
ser arg arg leu arg arg)OPA(arg arg arg   ile arg arg asn leu thr pro glu pro thr

181/61   SEQ ID NO: 638──◢   211/71
AGA CTG GCC CGC AGG TGA AGG CGG CGG ATT   CGG CGG AAT CTG ACG CCG GAG AGC TCG GCG
arg leu ala arg arg)OPA(arg arg arg ile   arg arg asn leu thr pro glu ser ser ala

241/81   SEQ ID NO: 639◢   271/91
AGG ACG CGT GCC CAG AAC AGG CCC TCG TCG   AGC GGC GCC CGT CGC GGT TGC GGC GAG GCT
arg thr arg ala gln asn arg pro ser ser   ser gly ala arg arg gly cys gly glu ala

301/101   331/111
GGC TTG TTG GCA TTG CGG CGA CGC TGC TCG   CGT TGG CCG GTG GCC TTG GCG CAG CGG GTT
gly leu leu ala leu arg arg arg cys ser   arg trp pro val ala leu ala gln arg val

361/121   391/131
ATT TTG CGT TGC GCT CAC ACC AGG AAA GCC   AAT CAA TCG CGC GCG AGG ACC TTG CGG CCA
ile leu arg cys ala his thr arg lys ala   asn gln ser arg ala arg thr leu arg pro

421/141   451/151
TTG AGG CCG CTA AGG ATT GCG TTG CGG CCA   CGC AGG CAC CCG ATG CTG GGG CGA TGT CGG
leu arg pro leu arg ile ala leu arg pro   arg arg his pro met leu gly arg cys arg

481/161
CTA GCA TGC AGA AGA TC
leu ala cys arg arg)

SEQ ID NOS:635-639

FIG. 38B

1/1 ———SEQ ID NO: 640   31/11
AGC GGT GAA CTG GTG GGC CCG GAT GGT TCA AGT ACG CCG TCG CAA ACT CGA GCA CAA CAG
(ser gly glu leu val gly pro asp gly ser ser thr pro ser gln thr arg ala gln gln
61/21 ———SEQ ID NO: 641   91/31
GAG ACG ACG GAT GGA AGG AGA TGC TGG CGC CGG CCA GCT GAA CCC TGC CGA TGC GAA TAA
glu thr thr asp gly arg arg cys trp arg arg pro ala glu pro cys arg cys glu)OCH
121/41   151/51
GTC GTC GTC TAC GGA GGT GAA GGC GGC GGA TTC GGC GGA ATC TGA CGC CGG AGC CGA CCA
(val val val tyr gly gly glu gly gly gly phe gly gly ile)OPA(arg arg ser arg pro
181/61 ———SEQ ID NO: 642   211/71   ———SEQ ID NO: 643
GAC TGG CCC GCA GGT GAA GGC GGC GGA TTC GGC GGA ATC TGA CGC CGG AGA GCT CGG CGA
asp trp pro ala gly glu gly gly gly phe gly gly ile)OPA(arg arg arg ala arg arg
241/81   271/91   ———SEQ ID NO: 644
GGA CGC GTG CCC AGA ACA GGC CCT CGT CGA GCG GCG CCC GTC GCG GTT GCG GCG AGG CTG
gly arg val pro arg thr gly pro arg arg ala ala pro val ala val ala ala arg leu
301/101   331/111
GCT TGT TGG CAT TGC GGC GAC GCT GCT CGC GTT GGC CGG TGG CCT TGG CGC AGC GGG TTA
ala cys trp his cys gly asp ala ala arg val gly arg trp pro trp arg ser gly leu
361/121   391/131
TTT TGC GTT GCG CTC ACA CCA GGA AAG CCA ATC AAT CGC GCG CGA GGA CCT TGC GGC CAT
phe cys val ala leu thr pro gly lys pro ile asn arg ala arg gly pro cys gly his)
421/141   451/151
TGA GGC CGC TAA GGA TTG CGT TGC GGC CAC GCA GGC ACC CGA TGC TGG GGC GAT GTC GGC
OPA gly arg OCH(gly leu arg cys gly his ala gly thr arg cys trp gly asp val gly)
481/161   ———SEQ ID NO: 645
TAG CAT GCA GAA GAT C
AMB(his ala glu asp)
     ———SEQ ID NO: 646

SEQ ID NOS:640-646

FIG. 38C

Sequence Rv0175 predicted by Cole et al., 1998 (Nature 393 537-544)
and containing seq38A

1/1 ___SEQ ID NO: 647   31/11
GTG AAG GCG GCG GAT TCG GCG GAA TCT GAC GCC GGA GCC GAC CAG ACT GGC CCG CAG GTG
(val lys ala ala asp ser ala glu ser asp ala gly ala asp gln thr gly pro gln val
61/21 ___SEQ ID NO: 648   91/31
AAG GCG GCG GAT TCG GCG GAA TCT GAC GCC GGA GAG CTC GGC GAG GAC GCG TGC CCA GAA
lys ala ala asp ser ala glu ser asp ala gly glu leu gly glu asp ala cys pro glu
121/41   151/51
CAG GCC CTC GTC GAG CGG CGC CCG TCG CGG TTG CGG CGA GGC TGG CTT GTT GGC ATT GCG
gln ala leu val glu arg arg pro ser arg leu arg arg gly trp leu val gly ile ala
181/61   211/71
GCG ACG CTG CTC GCG TTG GCC GGT GGC CTT GGC GCA GCG GGT TAT TTT GCG TTG CGC TCA
ala thr leu leu ala leu ala gly gly leu gly ala ala gly tyr phe ala leu arg ser
241/81   271/91
CAC CAG GAA AGC CAA TCA ATC GCG CGC GAG GAC CTT GCG GCC ATT GAG GCC GCT AAG GAT
his gln glu ser gln ser ile ala arg glu asp leu ala ala ile glu ala ala lys asp
301/101   331/111
TGC GTT GCG GCC ACG CAG GCA CCC GAT GCT GGG GCG ATG TCG GCT AGC ATG CAG AAG ATC
cys val ala ala thr gln ala pro asp ala gly ala met ser ala ser met gln lys ile
361/121   391/131
ATC GAG TGT GGC ACC GGT GAT TTC GGT GCC CAG GCG TCG TTG TAC ACC AGC ATG CTC GTC
ile glu cys gly thr gly asp phe gly ala gln ala ser leu tyr thr ser met leu val
421/141   451/151
GAG GCG TAT CAA GCG GCC AGC GTC CAC GTG CAA GTG ACC GAT ATG CGC GCG GCG GTC GAG
glu ala tyr gln ala ala ser val his val gln val thr asp met arg ala ala val glu
481/161   511/171
CGC AAC AAC AAT GAC GGG TCG GTC GAT GTT CTG GTG GCG CTC CGG GTC AAG GTG TCC AAC
arg asn asn asn asp gly ser val asp val leu val ala leu arg val lys val ser asn
541/181   571/191
ACC GAC TCG GAT GCC CAT GAA GTC GGC TAC CGT CTT CGG GTC CGG ATG GCA CTG GAT GAG
thr asp ser asp ala his glu val gly tyr arg leu arg val arg met ala leu asp glu
601/201   631/211
GGC CGC TAT AAG ATC GCC AAA CTC GAC CAG GTG ACG AAG TGA
gly arg tyr lys ile ala lys leu asp gln val thr lys)OPA

SEQ ID NOS:647-648

FIG. 38D

ORF according to Cole et al., 1998 (Nature 393 537-544) containing Rv0175

1/1 — SEQ ID NO: 649   31/11

```
TGA ACT GGT GGG GCC GGA TGG TGT CAA GTA CGC CGT CGC AAA CTC GAG CAC AAC AGG AGA
OPA(thr gly gly ala gly trp cys gln val arg arg arg lys leu glu his asn arg arg
```

61/21 — SEQ ID NO: 650   91/31

```
CGA CGG ATG GAA GGA GAT GCT GGC GCC GGC CAG CTG AAC CCT GCC GAT GCG AAT AAG TCG
arg arg met glu gly asp ala gly ala gly gln leu asn pro ala asp ala asn lys ser
```

121/41   151/51

```
TCG TCT ACG GAG GTG AAG GCG GCG GAT TCG GCG GAA TCT GAC GCC GGA GCC GAC CAG ACT
ser ser thr glu val lys ala ala asp ser ala glu ser asp ala gly ala asp gln thr
```

181/61   211/71

```
GGC CCG CAG GTG AAG GCG GCG GAT TCG GCG GAA TCT GAC GCC GGA GAG CTC GGC GAG GAC
gly pro gln val lys ala ala asp ser ala glu ser asp ala gly glu leu gly glu asp
```

241/81   271/91

```
GCG TGC CCA GAA CAG GCC CTC GTC GAG CGG CGC CCG TCG CGG TTG CGG CGA GGC TGG CTT
ala cys pro glu gln ala leu val glu arg arg pro ser arg leu arg arg gly trp leu
```

301/101   331/111

```
GTT GGC ATT GCG GCG ACG CTG CTC GCG TTG GCC GGT GGC CTT GGC GCA GCG GGT TAT TTT
val gly ile ala ala thr leu leu ala leu ala gly gly leu gly ala ala gly tyr phe
```

361/121   391/131

```
GCG TTG CGC TCA CAC CAG GAA AGC CAA TCA ATC GCG CGC GAG GAC CTT GCG GCC ATT GAG
ala leu arg ser his gln glu ser gln ser ile ala arg glu asp leu ala ala ile glu
```

421/141   451/151

```
GCC GCT AAG GAT TGC GTT GCG GCC ACG CAG GCA CCC GAT GCT GGG GCG ATG TCG GCT AGC
ala ala lys asp cys val ala ala thr gln ala pro asp ala gly ala met ser ala ser
```

481/161   511/171

```
ATG CAG AAG ATC ATC GAG TGT GGC ACC GGT GAT TTC GGT GCC CAG GCG TCG TTG TAC ACC
met gln lys ile ile glu cys gly thr gly asp phe gly ala gln ala ser leu tyr thr
```

541/181   571/191

```
AGC ATG CTC GTC GAG GCG TAT CAA GCG GCC AGC GTC CAC GTG CAA GTG ACC GAT ATG CGC
ser met leu val glu ala tyr gln ala ala ser val his val gln val thr asp met arg
```

601/201   631/211

```
GCG GCG GTC GAG CGC AAC AAC AAT GAC GGG TCG GTC GAT GTT CTG GTG GCG CTC CGG GTC
ala ala val glu arg asn asn asn asp gly ser val asp val leu val ala leu arg val
```

661/221   691/231

```
AAG GTG TCC AAC ACC GAC TCG GAT GCC CAT GAA GTC GGC TAC CGT CTT CGG GTC CGG ATG
lys val ser asn thr asp ser asp ala his glu val gly tyr arg leu arg val arg met
```

721/241   751/251

```
GCA CTG GAT GAG GGC CGC TAT AAG ATC GCC AAA CTC GAC CAG GTG ACG AAG TGA
ala leu asp glu gly arg tyr lys ile ala lys leu asp gln val thr lys)OPA
```

SEQ ID NOS:649-650

FIG. 38E

1/1 —SEQ ID NO: 651   31/11
ACA CCT CCC CCC CCG CCG CCG CTG CCG CCG GTT CCC TTT CCC AAG GAA TGT CCG GCG CCG
(thr pro pro pro pro pro pro leu pro pro val pro phe pro lys glu cys pro ala pro
61/21 —SEQ ID NO: 652   91/31
GGC GTG ATG CAA GGC TGC CTT GAG AGC ACC AGC GGC TTG ATC ATG GGC ATC GAC AGC AAG
gly val met gln gly cys leu glu ser thr ser gly leu ile met gly ile asp ser lys
121/41   151/51
ACC GCA CTG GTC GCC GAG CGC ATC ACC GGT GCC GTC GAG GAG ATC
thr ala leu val ala glu arg ile thr gly ala val glu glu ile)

SEQ ID NOS:651-652

FIG. 39A

1/1 —SEQ ID NO: 653   31/11
CAC CTC CCC CCC CGC CGC CGC TGC CGC CGG TTC CCT TTC CCA AGG AAT GTC CGG CGC CGG
(his leu pro pro arg arg arg cys arg arg phe pro phe pro arg asn val arg arg arg
61/21 —SEQ ID NO: 654   91/31
GCG TGA TGC AAG GCT GCC TTG AGA GCA CCA GCG GCT TGA TCA TGG GCA TCG ACA GCA AGA
ala)OPA(cys lys ala ala leu arg ala pro ala ala)OPA(ser trp ala ser thr ala arg
121/41 —SEQ ID NO: 655   151/51   —SEQ ID NO: 656
CCG CAC TGG TCG CCG AGC GCA TCA CCG GTG CCG TCG AGG AGA TC
pro his trp ser pro ser ala ser pro val pro ser arg arg)

SEQ ID NOS:653-656
FIG. 39B

1/1 —SEQ ID NO: 657   31/11
GAC ACC TCC CCC CCC GCC GCC GCT GCC GCC GGT TCC CTT TCC CAA GGA ATG TCC GGC GCC
(asp thr ser pro pro ala ala ala ala ala gly ser leu ser gln gly met ser gly ala
61/21 —SEQ ID NO: 658   91/31
GGG CGT GAT GCA AGG CTG CCT TGA GAG CAC CAG CGG CTT GAT CAT GGG CAT CGA CAG CAA
gly arg asp ala arg leu pro)OPA(glu his gln arg leu asp his gly his arg gln gln
121/41   SEQ NO: 659 —   151/51
GAC CGC ACT GGT CGC CGA GCG CAT CAC CGG TGC CGT CGA GGA GAT C
asp arg thr gly arg arg ala his his arg cys arg arg gly asp)

SEQ ID NOS:657-659

FIG. 39C

Coding sequence Rv3006 predicted by Cole et al., 1998 (Nature 393 537-544) and
containing seq39A

1/1 ——SEQ ID NO: 660   31/11
ATG TGG ACA ACG CGG TTG GTT CGA TCC GGA CTC GCC GCG CTG TGC GCG GCA GTG CTG GTA
(Met trp thr thr arg leu val arg ser gly leu ala ala leu cys ala ala val leu val
61/21 ——SEQ ID NO: 661   91/31
TCG AGC GGC TGC GCA CGG TTC AAC GAC GCT CAA TCT CAG CCG TTC ACC ACC GAA CCG GAG
ser ser gly cys ala arg phe asn asp ala gln ser gln pro phe thr thr glu pro glu
121/41   151/51
CTG CGG CCC CAA CCC AGC TCG ACA CCT CCC CCC CCG CCG CCG CTG CCG CCG GTT CCC TTT
leu arg pro gln pro ser ser thr pro pro pro pro pro pro leu pro pro val pro phe
181/61   211/71
CCC AAG GAA TGT CCG GCG CCG GGC GTG ATG CAA GGC TGC CTT GAG AGC ACC AGC GGC TTG
pro lys glu cys pro ala pro gly val met gln gly cys leu glu ser thr ser gly leu
241/81   271/91
ATC ATG GGC ATC GAC AGC AAG ACC GCA CTG GTC GCC GAG CGC ATC ACC GGT GCC GTC GAG
ile met gly ile asp ser lys thr ala leu val ala glu arg ile thr gly ala val glu
301/101   331/111
GAG ATC TCT ATC AGC GCC GAG CCG AAG GTA AAG ACG GTC ATC CCC GTG GAT CCT GCC GGT
glu ile ser ile ser ala glu pro lys val lys thr val ile pro val asp pro ala gly
361/121   391/131
GAC GGT GGC TTG ATG GAC ATT GTG CTG TCG CCC ACC TAC TCG CAA GAC CGG CTG ATG TAC
asp gly gly leu met asp ile val leu ser pro thr tyr ser gln asp arg leu met tyr
421/141   451/151
GCC TAC ATC AGC ACG CCC ACC GAC AAC CGG GTG GTG CGA GTG GCC GAC GGC GAC ATC CCC
ala tyr ile ser thr pro thr asp asn arg val val arg val ala asp gly asp ile pro
481/161   511/171
AAG GAC ATC CTG ACC GGC ATC CCC AAA GGT GCT GCC GGT AAC ACC GGG GCG CTG ATC TTC
lys asp ile leu thr gly ile pro lys gly ala ala gly asn thr gly ala leu ile phe
541/181   571/191
ACC AGT CCC ACC ACG CTG GTC GTG ATG ACC GGG GAT GCT GGC GAC CCG GCG TTG GCC GCC
thr ser pro thr thr leu val val met thr gly asp ala gly asp pro ala leu ala ala
601/201   631/211
GAT CCC CAA TCG TTG GCC GGT AAG GTC CTG CGT ATC GAA CAG CCC ACC ACC ATC GGC CAG
asp pro gln ser leu ala gly lys val leu arg ile glu gln pro thr thr ile gly gln
661/221   691/231
ACG CCG CCG ACG ACG GCG CTG TCT GGC ATC GGC TCC GGC GGC GGC TTG TGC ATC GAT CCG
thr pro pro thr thr ala leu ser gly ile gly ser gly gly gly leu cys ile asp pro
721/241   751/251
GTC GAC GGC TCG CTA TAT GTC GCC GAC CGC ACG CCA ACG GCG GAC CGA TTG CAG CGC ATC
val asp gly ser leu tyr val ala asp arg thr pro thr ala asp arg leu gln arg ile
781/261   811/271
ACC AAG AAC TCG GAG GTC TCT ACG GTA TGG ACC TGG CCG GAC AAG CCC GGC GTG GCC GGG
thr lys asn ser glu val ser thr val trp thr trp pro asp lys pro gly val ala gly
841/281   871/291
TGT GCC GCG ATG GAC GGC ACC GTG CTG GTC AAC CTG ATT AAT ACC AAA CTG ACG GTG GCG
cys ala ala met asp gly thr val leu val asn leu ile asn thr lys leu thr val ala
901/301   931/311
GTC CGG CTC GCG CCG TCG ACC GGT GCG GTC ACC GGA GAA CCC GAC GTT GTC CGC AAA GAC
val arg leu ala pro ser thr gly ala val thr gly glu pro asp val val arg lys asp
961/321   991/331
ACT CAT GCG CAT GCG TGG GCA TTA CGG ATG TCG CCG GAC GGC AAC GTC TGG GGA GCC ACC
thr his ala his ala trp ala leu arg met ser pro asp gly asn val trp gly ala thr
1021/341   1051/351
GTC AAC AAG ACC GCC GGC GAC GCC GAG AAG CTC GAC GAT GTG GTG TTC CCG CTG TTC CCG
val asn lys thr ala gly asp ala glu lys leu asp asp val val phe pro leu phe pro
1081/361   1111/371
CAG GGT GGC GGC TTC CCG CGC AAC AAC GAC GAC AAG ACC TGA
gln gly gly gly phe pro arg asn asn asp asp lys thr)OPA

SEQ ID NOS:660-661

FIG. 39D

ORF according to Cole et al., 1998 (Nature 393 537-544) and containing Rv3006

1/1 ⟋SEQ ID NO: 662   31/11
TAA GGC CAT TTA GTG CCG AAT TGG GGA TTT GAG CGG CGC TTT CGC CAG ACA ATC CGC ACA
OCH(gly his leu val pro asn trp gly phe glu arg arg phe arg gln thr ile arg thr
61/21 ⟋SEQ ID NO: 663   91/31
TTG ACC CTG ACC AGC CCA CCA AAA GGC CCC AAT TGG GCC GCC ATG CCG ACA GTG CGC ACC
leu thr leu thr ser pro pro lys gly pro asn trp ala ala met pro thr val arg thr
121/41   151/51
CCG GCA GGT GGC GGC GAT GCC CAC AAT GTC CGT AGC CTG TCG GTC ATG TGG ACA ACG CGG
pro ala gly gly gly asp ala his asn val arg ser leu ser val met trp thr thr arg
181/61   211/71
TTG GTT CGA TCC GGA CTC GCC GCG CTG TGC GCG GCA GTG CTG GTA TCG AGC GGC TGC GCA
leu val arg ser gly leu ala ala leu cys ala ala val leu val ser ser gly cys ala
241/81   271/91
CGG TTC AAC GAC GCT CAA TCT CAG CCG TTC ACC ACC GAA CCG GAG CTG CGG CCC CAA CCC
arg phe asn asp ala gln ser gln pro phe thr thr glu pro glu leu arg pro gln pro
301/101   331/111
AGC TCG ACA CCT CCC CCC CCG CCG CCG CTG CCG CCG GTT CCC TTT CCC AAG GAA TGT CCG
ser ser thr pro pro pro pro pro pro leu pro pro val pro phe pro lys glu cys pro
361/121   391/131
GCG CCG GGC GTG ATG CAA GGC TGC CTT GAG AGC ACC AGC GGC TTG ATC ATG GGC ATC GAC
ala pro gly val met gln gly cys leu glu ser thr ser gly leu ile met gly ile asp
421/141   451/151
AGC AAG ACC GCA CTG GTC GCC GAG CGC ATC ACC GGT GCC GTC GAG GAG ATC TCT ATC AGC
ser lys thr ala leu val ala glu arg ile thr gly ala val glu glu ile ser ile ser
481/161   511/171
GCC GAG CCG AAG GTA AAG ACG GTC ATC CCC GTG GAT CCT GCC GGT GAC GGT GGC TTG ATG
ala glu pro lys val lys thr val ile pro val asp pro ala gly asp gly gly leu met
541/181   571/191
GAC ATT GTG CTG TCG CCC ACC TAC TCG CAA GAC CGG CTG ATG TAC GCC TAC ATC AGC ACG
asp ile val leu ser pro thr tyr ser gln asp arg leu met tyr ala tyr ile ser thr
601/201   631/211
CCC ACC GAC AAC CGG GTG GTG CGA GTG GCC GAC GGC GAC ATC CCC AAG GAC ATC CTG ACC
pro thr asp asn arg val val arg val ala asp gly asp ile pro lys asp ile leu thr
661/221   691/231
GGC ATC CCC AAA GGT GCT GCC GGT AAC ACC GGG GCG CTG ATC TTC ACC AGT CCC ACC ACG
gly ile pro lys gly ala ala gly asn thr gly ala leu ile phe thr ser pro thr thr
721/241   751/251
CTG GTC GTG ATG ACC GGG GAT GCT GGC GAC CCG GCG TTG GCC GCC GAT CCC CAA TCG TTG
leu val val met thr gly asp ala gly asp pro ala leu ala ala asp pro gln ser leu
781/261   811/271
GCC GGT AAG GTC CTG CGT ATC GAA CAG CCC ACC ACC ATC GGC CAG ACG CCG CCG ACG ACG
ala gly lys val leu arg ile glu gln pro thr thr ile gly gln thr pro pro thr thr
841/281   871/291
GCG CTG TCT GGC ATC GGC TCC GGC GGC GGC TTG TGC ATC GAT CCG GTC GAC GGC TCG CTA
ala leu ser gly ile gly ser gly gly gly leu cys ile asp pro val asp gly ser leu
901/301   931/311
TAT GTC GCC GAC CGC ACG CCA ACG GCG GAC CGA TTG CAG CGC ATC ACC AAG AAC TCG GAG
tyr val ala asp arg thr pro thr ala asp arg leu gln arg ile thr lys asn ser glu

SEQ ID NOS:662-663

FIG. 39E

961/321   991/331
GTC TCT ACG GTA TGG ACC TGG CCG GAC AAG   CCC GGC GTG GCC GGG TGT GCC GCG ATG GAC
val ser thr val trp thr trp pro asp lys   pro gly val ala gly cys ala ala met asp
1021/341   1051/351
GGC ACC GTG CTG GTC AAC CTG ATT AAT ACC   AAA CTG ACG GTG GCG GTC CGG CTC GCG CCG
gly thr val leu val asn leu ile asn thr   lys leu thr val ala val arg leu ala pro
1081/361   1111/371
TCG ACC GGT GCG GTC ACC GGA GAA CCC GAC   GTT GTC CGC AAA GAC ACT CAT GCG CAT GCG
ser thr gly ala val thr gly glu pro asp   val val arg lys asp thr his ala his ala
1141/381   1171/391
TGG GCA TTA CGG ATG TCG CCG GAC GGC AAC   GTC TGG GGA GCC ACC GTC AAC AAG ACC GCC
trp ala leu arg met ser pro asp gly asn   val trp gly ala thr val asn lys thr ala
1201/401   1231/411
GGC GAC GCC GAG AAG CTC GAC GAT GTG GTG   TTC CCG CTG TTC CCG CAG GGT GGC GGC TTC
gly asp ala glu lys leu asp asp val val   phe pro leu phe pro gln gly gly gly phe
1261/421
CCG CGC AAC AAC GAC GAC AAG ACC TGA
pro arg asn asn asp asp lys thr)OPA

SEQ ID NOS:662-663 (continued)

FIG. 39E (continued)

1/1 ⟋SEQ ID NO: 664   31/11
GAA GGC CTT GTT GAG CCG GCG CAC GAA AAC   GAT CGT TGT GTG TAC ATT GGT GTG TAT GGC
(glu gly leu val glu pro ala his glu asn   asp arg cys val tyr ile gly val tyr gly
61/21 ⌐SEQ ID NO: 665   91/31
TCG GTT GAA CGT GTA TGT GCC CGA CGA ATT   GGC GGA GCG CGC CAG GGC GCG GGG CTT GAA
ser val glu arg val cys ala arg arg ile   gly gly ala arg gln gly ala gly leu glu
121/41   151/51
CGT CTC GGC GCT GAC TCA GGC CGC GAT CAG   TGC CGA GTT GGA GAA CTC CGC AAC CGA TGC
arg leu gly ala asp ser gly arg asp gln   cys arg val gly glu leu arg asn arg cys
181/61   211/71
GTG GCT TGA GGG GTT GGA ACC CAG AAG CAC   CGG CGC TCG GCA TGA TGA CGT GCT GGG TGC
val ala)OPA(gly val gly thr gln lys his   arg arg ser ala)OPA OPA(arg ala gly cys
241/81   ⌐SEQ ID NO: 931   271/91   ⌐SEQ ID NO: 932
GAT CGA TGC CGC TCG CGA TGA GTT CGA AGC   GTG AGA GCA TCG CCC ACT TCG CCG CCG GAG
asp arg cys arg ser arg)OPA(val arg ser   val arg ala ser pro thr ser pro pro glu
301/101   SEQ ID NO: 933⌐   331/111
CAG GTG GTC GTC GAC GCG AGT GCC ATG GTG GAT C
gln val val val asp ala ser ala met val asp)

SEQ ID NOS:664-665, 931-933

FIG. 40A

1/1 ⸺SEQ ID NO: 666   31/11
AAG GCC TTG TTG AGC CGG CGC ACG AAA ACG ATC GTT GTG TGT ACA TTG GTG TGT ATG GCT
(lys ala leu leu ser arg arg thr lys thr ile val val cys thr leu val cys met ala
61/21 ⸺SEQ ID NO: 667   91/31
CGG TTG AAC GTG TAT GTG CCC GAC GAA TTG GCG GAG CGC GCC AGG GCG CGG GGC TTG AAC
arg leu asn val tyr val pro asp glu leu ala glu arg ala arg ala arg gly leu asn
121/41   151/51
GTC TCG GCG CTG ACT CAG GCC GCG ATC AGT GCC GAG TTG GAG AAC TCC GCA ACC GAT GCG
val ser ala leu thr gln ala ala ile ser ala glu leu glu asn ser ala thr asp ala
181/61   211/71
TGG CTT GAG GGG TTG GAA CCC AGA AGC ACC GGC GCT CGG CAT GAT GAC GTG CTG GGT GCG
trp leu glu gly leu glu pro arg ser thr gly ala arg his asp asp val leu gly ala
241/81   271/91
ATC GAT GCC GCT CGC GAT GAG TTC GAA GCG TGA GAG CAT CGC CCA CTT CGC CGC CGG AGC
ile asp ala ala arg asp glu phe glu ala)OPA(glu his arg pro leu arg arg arg ser
301/101   331/111 ⸺SEQ ID NO: 668
AGG TGG TCG TCG ACG CGA GTG CCA TGG TGG ATC
arg trp ser ser thr arg val pro trp trp ile)

SEQ ID NOS:666-668

FIG. 40B

1/1 ⸺SEQ ID NO: 669   31/11
AGG CCT TGT TGA GCC GGC GCA CGA AAA CGA TCG TTG TGT GTA CAT TGG TGT GTA TGG CTC
arg pro cys OPA(ala gly ala arg lys arg ser leu cys val his trp cys val trp leu
61/21   SEQ ID NO: 670   91/31
GGT TGA ACG TGT ATG TGC CCG ACG AAT TGG CGG AGC GCG CCA GGG CGC GGG GCT TGA ACG
gly)OPA(thr cys met cys pro thr asn trp arg ser ala pro gly arg gly ala)OPA(thr
121/41   ⸺SEQ ID NO: 671   151/51   SEQ ID NO: 672 ⸺
TCT CGG CGC TGA CTC AGG CCG CGA TCA GTG CCG AGT GGA GAA CTC CGC AAC CGA TGC GT
ser arg arg)OPA(leu arg pro arg ser val pro ser trp arg thr pro gln pro met arg
181/61   ⸺SEQ ID NO: 673   211/71
GGC TTG AGG GGT TGG AAC CCA GAA GCA CCG GCG CTC GGC ATG ATG ACG TGC TGG GTG CGA
gly leu arg gly trp asn pro glu ala pro ala leu gly met met thr cys trp val arg
241/81   271/91
TCG ATG CCG CTC GCG ATG AGT TCG AAG CGT GAG AGC ATC GCC CAC TTC GCC GCC GGA GCA
ser met pro leu ala met ser ser lys arg glu ser ile ala his phe ala ala gly ala
301/101
GGT GGT CGT CGA CGC GAG TGC CAT GGT GGA TC
gly gly arg arg arg glu cys his gly gly)

SEQ ID NOS:669-673

FIG. 40C

Coding sequence Rv0549c predicted by Cole et al., 1998 (Nature 393:537-544) and containing seq40A

```
1/1    SEQ ID NO: 674                    31/11
gtg aga gca tcg ccc act tcg ccg ccg gag cag gtg gtc gtc gac gcg agt gcc atg gtg
(val arg ala ser pro thr ser pro pro glu gln val val val asp ala ser ala met val
61/21   SEQ ID NO: 675                   91/31
gat cta ctg gct cgc act agc gat cgg tgc tct gcg gtg cgc gcg cgg ctg gct cgg acc
asp leu leu ala arg thr ser asp arg cys ser ala val arg ala arg leu ala arg thr
121/41                                   151/51
gcg atg cac gcg ccg gcg cac ttc gat gca gag gtg ttg tcg gcg ctg ggg cgc atg cag
ala met his ala pro ala his phe asp ala glu val leu ser ala leu gly arg met gln
181/61                                   211/71
cgc gcc ggc gca ctc acc gtt gcc tat gtc gat gcg gca ctg gag gag ttg cga cag gtg
arg ala gly ala leu thr val ala tyr val asp ala ala leu glu glu leu arg gln val
241/81                                   271/91
ccg gtg act cga cac ggt ctt tcg tcg ctg ctt gct gga gcg tgg tcg cgc cgc gac acc
pro val thr arg his gly leu ser ser leu leu ala gly ala trp ser arg arg asp thr
301/101                                  331/111
ctc cgc ctg acc gat gcc ctc tac gtc gag ctg gcc gaa acg gca ggt ctg gtg ttg ttg
leu arg leu thr asp ala leu tyr val glu leu ala glu thr ala gly leu val leu leu
361/121                                  391/131
acc acc gac gaa aga ttg gca cgc gcc tgg ccc tcg gct cac gcc atc ggc tga
thr thr asp glu arg leu ala arg ala trp pro ser ala his ala ile gly)OPA
```

SEQ ID NOS:674-675

FIG. 40D

ORF according to Cole et al., 1998 (Nature 393:537-544) and containing Rv0549c

```
1/1    SEQ ID NO: 676                    31/11
tga gtt cga agc gtg aga gca tcg ccc act tcg ccg ccg gag cag gtg gtc gtc gac gcg
OPA(val arg ser val arg ala ser pro thr ser pro pro glu gln val val val asp ala
61/21   SEQ ID NO: 677                   91/31
agt gcc atg gtg gat cta ctg gct cgc act agc gat cgg tgc tct gcg gtg cgc gcg cgg
ser ala met val asp leu leu ala arg thr ser asp arg cys ser ala val arg ala arg
121/41                                   151/51
ctg gct cgg acc gcg atg cac gcg ccg gcg cac ttc gat gca gag gtg ttg tcg gcg ctg
leu ala arg thr ala met his ala pro ala his phe asp ala glu val leu ser ala leu
181/61                                   211/71
ggg cgc atg cag cgc gcc ggc gca ctc acc gtt gcc tat gtc gat gcg gca ctg gag gag
gly arg met gln arg ala gly ala leu thr val ala tyr val asp ala ala leu glu glu
241/81                                   271/91
ttg cga cag gtg ccg gtg act cga cac ggt ctt tcg tcg ctg ctt gct gga gcg tgg tcg
leu arg gln val pro val thr arg his gly leu ser ser leu leu ala gly ala trp ser
301/101                                  331/111
cgc cgc gac acc ctc cgc ctg acc gat gcc ctc tac gtc gag ctg gcc gaa acg gca ggt
arg arg asp thr leu arg leu thr asp ala leu tyr val glu leu ala glu thr ala gly
361/121                                  391/131
ctg gtg ttg ttg acc acc gac gaa aga ttg gca cgc gcc tgg ccc tcg gct cac gcc atc
leu val leu leu thr thr asp glu arg leu ala arg ala trp pro ser ala his ala ile
421/141
ggc tga
gly)OPA
```

SEQ ID NOS:676-677

FIG. 40E

1/1 ────── SEQ ID NO:678   31/11
CCT GGC CGG GAC GCC TAC GTG TAG CCC GCG GCT AGC ACA GGA TAG CCA TTG TTG TGC GGT
(pro gly arg asp ala tyr val)AMB(pro ala ala ser thr gly)AMB(pro leu leu cys gly
61/21 ──SEQ ID NO: 679   SEQ ID NO: 680 91/31   ──SEQ ID NO: 681
AGC GCC AAA ACG ATC AGC CCT TCG CGG ACA TGT CAG CAC CCG CCT TGG CCG GGA GAG CGG
ser ala lys thr ile ser pro ser arg thr cys gln his pro pro trp pro gly glu arg
121/41   151/51
CGT CGT GAC CGT GCT GTC ACC ACG TCT GGT TAG GCT CGG GGC GCG GGC TGG CGC GGA GGA
arg arg asp arg ala val thr thr ser gly)AMB(ala arg gly ala gly trp arg gly gly
181/61   211/71 ──SEQ ID NO: 682
GGT GTG TTG CGG AGG AGG TGT GTT GTA GTG GGG ACG GCG GAT CGG CCG TTG GAC GCC TCG
gly val leu arg arg arg cys val val val gly thr ala asp arg pro leu asp ala ser
241/81   271/91
GCC TTG CGG GAC TGG GCA CAC GCC GTC GTC AGC GAT C
ala leu arg asp trp ala his ala val val ser asp)

SEQ ID NOS:678-682

FIG. 41A

1/1 ──SEQ ID NO: 683   31/11
CTG GCC GGG ACG CCT ACG TGT AGC CCG CGG CTA GCA CAG GAT AGC CAT TGT TGT GCG GTA
(leu ala gly thr pro thr cys ser pro arg leu ala gln asp ser his cys cys ala val
61/21 ──SEQ ID NO: 684   91/31
GCG CCA AAA CGA TCA GCC CTT CGC GGA CAT GTC AGC ACC CGC CTT GGC CGG GAG AGC GGC
ala pro lys arg ser ala leu arg gly his val ser thr arg leu gly arg glu ser gly
121/41   151/51
GTC GTG ACC GTG CTG TCA CCA CGT CTG GTT AGG CTC GGG GCG CGG GCT GGC GCG GAG GAG
val val thr val leu ser pro arg leu val arg leu gly ala arg ala gly ala glu glu
181/61   211/71
GTG TGT TGC GGA GGA GGT GTG TTG TAG TGG GGA CGG CGG ATC GGC CGT TGG ACG CCT CGG
val cys cys gly gly gly val leu)AMB(trp gly arg arg ile gly arg trp thr pro arg
241/81   SEQ ID NO: 685 271/91
CCT TGC GGG ACT GGG CAC ACG CCG TCG TCA GCG ATC
pro cys gly thr gly his thr pro ser ser ala ile)

SEQ ID NOS:683-685

FIG. 41B

1/1 SEQ ID NO: 686   31/11
TGG CCG GGA CGC CTA CGT GTA GCC CGC GGC TAG CAC AGG ATA GCC ATT GTT GTG CGG TAG
(trp pro gly arg leu arg val ala arg gly)AMB(his arg ile ala ile val val arg)AMB
61/21 SEQ ID NO: 687   91/31 SEQ ID NO: 688
CGC CAA AAC GAT CAG CCC TTC GCG GAC ATG TCA GCA CCC GCC TTG GCC GGG AGA GCG GCG
(arg gln asn asp gln pro phe ala asp met ser ala pro ala leu ala gly arg ala ala
121/41 SEQ ID NO: 689   151/51
TCG TGA CCG TGC TGT CAC CAC GTC TGG TTA GGC TCG GGG CGC GGG CTG GCG CGG AGG AGG
ser)OPA(pro cys cys his his val trp leu gly ser gly arg gly leu ala arg arg arg
181/61 SEQ ID NO: 690   211/71
TGT GTT GCG GAG GAG GTG TGT TGT AGT GGG GAC GGC GGA TCG GCC GTT GGA CGC CTC GGC
cys val ala glu glu val cys cys ser gly asp gly gly ser ala val gly arg leu gly
241/81   271/91
CTT GCG GGA CTG GGC ACA CGC CGT CGT CAG CGA TC
leu ala gly leu gly thr arg arg arg gln arg)

SEQ ID NOS:686-690

FIG. 41C

Coding sequence Rv2975c predicted by Cole et al, 1998 (Nature 393: 537-544) and containing seq41A

1/1 SEQ ID NO: 691   31/11
gtg ggg acg gcg gat cgg ccg ttg gac gcc tcg gcc ttg cgg gac tgg gca cac gcc gtc
(val gly thr ala asp arg pro leu asp ala ser ala leu arg asp trp ala his ala val
61/21 SEQ ID NO: 692   91/31
gtc agc gat ctg atc ctc cac atc gac gag atc aac cgg ctc aat gtg ttc ccg gtc gct
val ser asp leu ile leu his ile asp glu ile asn arg leu asn val phe pro val ala
121/41   151/51
gac tcc gat acc ggc gtc aac atg ctg ttc acc atg cgt gcc gcg gtc gta gaa gct gat
asp ser asp thr gly val asn met leu phe thr met arg ala ala val val glu ala asp
181/61   211/71
ttg cac gcg aat tcg cag gct gac gcc gaa gac gtg gcg cgg gtt gcg gcc gct ctc gcg
leu his ala asn ser gln ala asp ala glu asp val ala arg val ala ala ala leu ala
241/81
gcc ggc gcg cgt tga
ala gly ala arg)OPA

SEQ ID NOS:691-692

FIG. 41D

ORF according to Cole et al, 1998 (Nature 393: 537-544) and containing Rv2975c

```
1/1 ___SEQ ID NO: 693                          31/11
tag gct cgg ggc gcg ggc tgg cgc gga gga ggt gtg ttg cgg agg agg tgt gtt gta gtg
AMB(ala arg gly ala gly trp arg gly gly gly val leu arg arg arg cys val val val
61/21 ‾‾‾SEQ ID NO: 694                         91/31
ggg acg gcg gat cgg ccg ttg gac gcc tcg gcc ttg cgg gac tgg gca cac gcc gtc gtc
gly thr ala asp arg pro leu asp ala ser ala leu arg asp trp ala his ala val val
121/41                                          151/51
agc gat ctg atc ctc cac atc gac gag atc aac cgg ctc aat gtg ttc ccg gtc gct gac
ser asp leu ile leu his ile asp glu ile asn arg leu asn val phe pro val ala asp
181/61                                          211/71
tcc gat acc ggc gtc aac atg ctg ttc acc atg cgt gcc gcg gtc gta gaa gct gat ttg
ser asp thr gly val asn met leu phe thr met arg ala ala val val glu ala asp leu
241/81                                          271/91
cac gcg aat tcg cag gct gac gcc gaa gac gtg gcg cgg gtt gcg gcc gct ctc gcg gcc
his ala asn ser gln ala asp ala glu asp val ala arg val ala ala ala leu ala ala
301/101
ggc gcg cgt tga
gly ala arg)OPA
```

SEQ ID NOS:693-694

FIG. 41E

sequence Rv 2974C predicted by Cole et al. (Nature 393:537-544) and which may be in the same reading frame as Seq41D. The sequencing of this region reveals, in one case out of three, a deletion of two nucleotides putting in phase observed in

```
1/1 ___SEQ ID NO: 695                          31/11
ttg aac gga gct cgc ggc aac tcc ggc gtg atc ctg tcc cag atc ctg cgc ggg atc gca
(leu asn gly ala arg gly asn ser gly val ile leu ser gln ile leu arg gly ile ala
61/21 ‾‾‾SEQ ID NO: 696                         91/31
gag gtg acc gcg act gcg gcc gcc gcc tct ggc gcg gta ttg cgg gcg gtc gac gcc aac
glu val thr ala thr ala ala ala ala ser gly ala val leu arg ala val asp ala asn
121/41                                          151/51
gcc ctc ggg gcc gcg ttg tgg cgc ggc gtc gag ttg gtc gtc gcg tcg atg ggt ggc gtg
ala leu gly ala ala leu trp arg gly val glu leu val val ala ser met gly gly val
181/61                                          211/71
gag gtg ccg gga act atc gtc tcg gtg ctg cgg gcc gcc gcc gga gcc gtc gac cag tgc
glu val pro gly thr ile val ser val leu arg ala ala ala gly ala val asp gln cys
241/81                                          271/91
gcg cac gag ggg ttg gcc ggt gcg gtc acc gcc gcc ggt gac gcg gcg gtc atc gcg ctg
ala his glu gly leu ala gly ala val thr ala ala gly asp ala ala val ile ala leu
301/101                                         331/111
gaa aag acc ccc gaa cag ctt gac gtg ctc gcc gat gcg ggc gcg gtg gac gcc ggc gga
glu lys thr pro glu gln leu asp val leu ala asp ala gly ala val asp ala gly gly
```

SEQ ID NOS:695-696

FIG. 41F

361/121
cgg ggc ctg ctg gtt ctg ctg gac gcg ttg
arg gly leu leu val leu leu asp ala leu
421/141
cgg gcg gtc tac gaa ccc tcg ccg cgc gcg
arg ala val tyr glu pro ser pro arg ala
481/161
gcc ccg caa ttc gag gtg atg tat ctg ttg
ala pro gln phe glu val met tyr leu leu
541/181
ttg cgg gat cga ctc aag gaa ttg ggt gag
leu arg asp arg leu lys glu leu gly glu
601/201
agc tac tcc gta cac gtc cac acc gac gac
ser tyr ser val his val his thr asp asp
661/221
gtg ggg cga gtt agc cgg atc gtg atc tcg
val gly arg val ser arg ile val ile ser
721/241
gcc ggt ggc tgg acg cgg ggc cgc gcc gtg
ala gly gly trp thr arg gly arg ala val
781/261
gag ctg ttc gcc ggg gag ggc gcc tgc gtg
glu leu phe ala gly glu gly ala cys val
841/281
gcc gcc gat atc agt gcc cac cag ctg gtg
ala ala asp ile ser ala his gln leu val
901/301
gtg atg gtg ctg ccc aat ggc tat gtg gcc
val met val leu pro asn gly tyr val ala
961/321
gcg atc ggc tgg ggc gtc gac gtg gta ccc
ala ile gly trp gly val asp val val pro
1021/341
gcc gcg ctg gcc gtg cat gac gcg gcc cgc
ala ala leu ala val his asp ala ala arg
1081/361
cgt gcc gcc ggt gct tcc cgg cac gga tcg
arg ala ala gly ala ser arg his gly ser
1141/381
tgg gcc ggt acc tgc aag ccg ggc gac ggt
trp ala gly thr cys lys pro gly asp gly
1201/401
gtc gcc gac gat gtc gcc gcg gcg gcc atc
val ala asp asp val ala ala ala ala ile
1261/421
ggc gat ctg gtg acg gtg cta att ggc gcc
gly asp leu val thr val leu ile gly ala
1321/441
gaa cgg cat gtg cac gac cac cat cca ggc
glu arg his val his asp his his pro gly
1381/461
cgc ggc gac gcg ctg ctg atc ggg gtc gag tag
arg gly asp ala leu leu ile gly val glu)AMB

391/131
cgc tcc acc atc tgc ggg cag gca cct gcc
arg ser thr ile cys gly gln ala pro ala
451/151
ttg ccg acc gac acg gct acc caa cgc ccc
leu pro thr asp thr ala thr gln arg pro
511/171
gcg gta tgt gat gct gca gcg gcg gac cag
ala val cys asp ala ala ala ala asp gln
571/191
tcg gtg gcc atc gcc gct gct ccg ccc gac
ser val ala ile ala ala ala pro pro asp
631/211
gcc ggt gcc gcc gtg gaa gcc gga ttg gcg
ala gly ala ala val glu ala gly leu ala
691/231
gcg ctc ggt tcc ggg acc agc gga ttg ccg
ala leu gly ser gly thr ser gly leu pro
751/251
ctg gcg gtc gtc gac ggc gac ggt gcc gcc
leu ala val val asp gly asp gly ala ala
811/271
ctg cga ccg ggt cca gac gcc gtg aca ccg
leu arg pro gly pro asp ala val thr pro
871/291
cgg gcc gtg gta gac acc ggc gcc gcg cac
arg ala val val asp thr gly ala ala his
931/311
gcc gaa gaa ctg gtg gcc ggg tgt acc gcg
ala glu glu leu val ala gly cys thr ala
991/331
gtg ccg acc gga tcg atg gtg cag ggg ttg
val pro thr gly ser met val gln gly leu
1051/351
cag gcc gtc gac gac ggc tac agc atg gcc
gln ala val asp asp gly tyr ser met ala
1111/371
gtg cgc att gcc acc caa aag gcg ctg acc
val arg ile ala thr gln lys ala leu thr
1171/391
ctg ggt atc gcg ggc gac gag gtg ctg atc
leu gly ile ala gly asp glu val leu ile
1231/411
ggt ctg gtc gac ctg ttg ttg gca tcg gga
gly leu val asp leu leu leu ala ser gly
1291/431
ggc gta acc gaa gac gtg gct gtc gtc ctg
gly val thr glu asp val ala val val leu
1351/451
ggc acc gag ctg gtc tcc tac cgc acc gga cac
gly thr glu leu val ser tyr arg thr gly his
1411/471

SEQ ID NOS:695-696 (continued)

FIG. 41F (continued)

SEQ ID NO: 697: first reading frame nucleic acid
SEQ ID NO: 702: second reading frame nucleic acid
SEQ ID NO: 717: third reading frame nucleic acid
Seq41T comprising seq 41F and seq 41S ⟶ SEQ ID NO: 698
1/1 ⟋ SEQ ID NO: 697,702,717   31/11
tta ggc tcg ggg cgc ggg ctg/gcg cgg agg agg tgt gtt gcg gag gag gtg tgt tgt agt
leu gly ser gly arg gly leu ala arg arg arg cys val ala glu glu val cys cys ser
 AMB(ala arg gly ala gly trp arg gly gly gly val leu arg arg arg cys val val val
  (arg\leu gly ala arg ala gly ala glu glu val cys cys gly gly gly val leu)AMB(trp
61/21 SEQ ID NO: 703 ⟋ SEQ ID NO: 718   91/31   SEQ ID NO: 719 ⟶
ggg gac ggc gga tcg gcc gtt gga cgc ctc ggc ctt gcg gga ctg ggc aca cgc cgt cgt
gly asp gly gly ser ala val gly arg leu gly leu ala gly leu gly thr arg arg arg
 gly thr ala asp arg pro leu asp ala ser ala leu arg asp trp ala his ala val val
  gly arg ile gly arg trp thr pro arg pro cys gly thr gly his thr pro ser ser
121/41   151/51
cag cga tct gat cct cca cat cga cga gat caa ccg gct caa tgt gtt ccc ggt cgc tga
gln arg ser asp pro pro his arg arg asp gln pro ala gln cys val pro gly arg)OPA
 ser asp leu ile leu his ile asp glu ile asn arg leu asn val phe pro val ala asp
  ala ile)OPA(ser ser thr ser thr arg ser thr gly ser met cys ser arg ser leu thr
181/61   ⟍SEQ ID NO: 720   211/71
ctc cga tac cgg cgt caa cat gct gtt cac cat gcg tgc cgc ggt cgt aga agc tga ttt
(leu arg tyr arg arg gln his ala val his his ala cys arg gly arg arg arg ser)OPA(phe
  ser\asp thr gly val asn met leu phe thr met arg ala ala val val glu ala asp(leu
   pro\ile pro ala ser thr cys cys ser pro cys val pro arg ser)AMB(lys leu ile)cys
241/81 SEQ ID NO: 699   271/91 SEQ ID NO: 721 ⟋SEQ ID NO: 700
gca cgc gaa ttc gca ggc tga cgc cga aga cgt ggc gcg ggt tgc ggc cgc tct cgc ggc
ala arg glu phe ala gly)OPA(arg arg arg arg gly ala gly cys gly arg ser arg gly
 his ala asn ser gln ala asp\ala glu asp val ala arg val ala ala ala leu ala ala
  thr arg ile arg arg leu thr) pro lys thr trp arg gly leu arg pro leu ser arg pro
301/101   SEQ ID NO: 701 ⟶   331/111
cgg cgc gcg ttg aac gga gct cgc ggc aac tcc ggc gtg atc ctg tcc cag atc ctg cgc
arg arg ala leu asn gly ala arg gly asn ser gly val ile leu ser gln ile leu arg
 gly ala arg)OPA(thr glu leu ala ala thr pro ala)OPA(ser cys pro arg ser cys ala
  ala arg val glu\arg ser ser arg gln leu arg arg asp\pro val pro asp pro ala arg
361/121   ⟍SEQ ID NO: 704   391/131   ⟍SEQ ID NO: 705
ggg atc gca gag gtg acc gcg act gcg gcc gcc gcc tct ggc gcg gta ttg cgg gcg gtc
gly ile ala glu val thr ala thr ala ala ala ala ser gly ala val leu arg ala val
 gly ser gln arg)OPA(pro arg leu arg pro pro pro leu ala arg tyr cys gly arg ser
  asp arg arg gly asp\arg asp cys gly arg arg leu trp arg gly ile ala gly gly arg
421/141   ⟍SEQ ID NO: 706 451/151
gac gcc aac gcc ctc ggg gcc gcg ttg tgg cgc ggc gtg gag ttg gtc gtc gcg tcg atg
asp ala asn ala leu gly ala ala leu trp arg gly val glu leu val val ala ser met
 thr pro thr pro ser gly pro arg cys gly ala ala ser ser trp ser ser arg arg trp
  arg gln arg pro arg gly arg val val ala arg arg arg val gly arg arg val asp gly
481/161   511/171
ggt ggc gtg gag gtg ccg gga act atc gtc tcg gtg ctg cgg gcc gcc gcc gga gcc gtc
gly gly val glu val pro gly thr ile val ser val leu arg ala ala ala gly ala val
 val ala trp arg cys arg glu leu ser ser arg cys cys gly pro pro pro glu pro ser
  trp arg gly gly ala gly asn tyr arg leu gly ala ala gly arg arg arg ser arg arg
541/181   571/191
gac cag tgc gcg cac gag ggg ttg gcc ggt gcg gtc acc gcc gcc ggt gac gcg gcg gtc
asp gln cys ala his glu gly leu ala gly ala val thr ala ala gly asp ala ala val
 thr ser ala arg thr arg gly trp pro val arg ser pro pro pro val thr arg arg ser
  pro val arg ala arg gly val gly arg cys gly his arg arg arg)OPA(arg gly gly his
601/201   631/211 _ SEQ ID NO: 722 ⟋
atc gcg ctg gaa aag acc ccc gaa cag ctt gac gtg ctc gcc gat gcg ggc gcg gtg gac
ile ala leu glu lys thr pro glu gln leu asp val leu ala asp ala gly ala val asp
 ser arg trp lys arg pro pro asn ser leu thr cys ser pro met arg ala arg trp thr
  arg ala gly lys asp pro arg thr ala)OPA(arg ala arg arg cys gly arg gly gly arg
661/221   691/231 ⟍SEQ ID NO: 723
gcc ggc gga cgg ggc ctg ctg gtt ctg ctg gac gcg ttg cgc tcc acc atc tgc ggg cag
ala gly gly arg gly leu leu val leu leu asp ala leu arg ser thr ile cys gly gln
 pro ala asp gly ala cys trp phe cys trp thr arg cys ala pro pro ser ala gly arg
  arg arg thr gly pro ala gly ser ala gly arg val ala leu his his leu arg ala gly

SEQ ID NOS:697-727

FIG. 41G

```
721/241                                 751/251
gca cct gcc cgg gcg gtc tac gaa ccc tcg ccg cgc gcg ttg ccg acc gac acg gct acc
ala pro ala arg ala val tyr glu pro ser pro arg ala leu pro thr asp thr ala thr
  his leu pro gly arg ser thr asn pro arg arg ala arg cys arg pro thr arg leu pro
    thr cys pro gly gly leu arg thr leu ala ala arg val ala asp arg his gly tyr pro
781/261                                 811/271
caa cgc ccc gcc ccg caa ttc gag gtg atg tat ctg ttg gcg gta tgt gat gct gca gcg
gln arg pro ala pro gln phe glu val met tyr leu leu ala val cys asp ala ala ala
  asn ala pro pro arg asn ser arg)OPA(cys ile cys trp arg tyr val met leu gln arg
    thr pro arg pro ala ile arg gly/asp val ser val gly gly met)OPA(cys cys ser gly
841/281        SEQ ID NO: 707               871/291  SEQ ID NO: 724
gcg gac cag ttg cgg gat cga ctc aag gaa ttg ggt gag tcg gtg gcc atc gcc gct gct
ala asp gln leu arg asp arg leu lys glu leu gly glu ser val ala ile ala ala ala
  arg thr ser cys gly ile asp ser arg asn trp val ser arg trp pro ser pro leu leu
    gly pro val ala gly ser thr gln gly ile gly)OPA(val gly gly his arg arg cys ser
901/301                                 931/311    SEQ ID NO: 725
ccg ccc gac agc tac tcc gta cac gtc cac acc gac gac gcc ggt gcc gcc gtg gaa gcc
pro pro asp ser tyr ser val his val his thr asp asp ala gly ala ala val glu ala
  arg pro thr ala thr pro tyr thr ser thr pro thr thr pro val pro pro trp lys pro
    ala arg gln leu leu arg thr arg pro his arg arg arg arg cys arg arg gly ser arg
961/321                                 991/331
gga ttg gcg gtg ggg cga gtt agc cgg atc gtg atc tcg gcg ctc ggt tcc ggg acc agc
gly leu ala val gly arg val ser arg ile val ile ser ala leu gly ser gly thr ser
  asp trp arg trp gly glu leu ala gly ser)OPA(ser arg arg ser val pro gly pro ala
    ile gly gly gly ala ser)AMB(pro asp arg asp leu gly ala arg phe arg asp gln arg
1021/341      SEQ ID NO: 726              1051/351       SEQ ID NO: 708
gga ttg ccg gcc ggt ggc tgg acg cgg ggc cgc gcc gtg ctg gcg gtc gtc gac ggc gac
gly leu pro ala gly gly trp thr arg gly arg ala val leu ala val val asp gly asp
  asp cys arg pro val ala gly arg gly ala ala pro cys trp arg ser ser thr ala thr
    ile ala gly arg trp leu asp ala gly pro arg arg ala gly gly arg arg arg arg arg
1081/361                                1111/371
ggt gcc gcc gag ctg ttc gcc ggg gag ggc gcc tgc gtg ctg cga ccg ggt cca gac gcc
gly ala ala glu leu phe ala gly glu gly ala cys val leu arg pro gly pro asp ala
  val pro pro ser cys ser pro gly arg ala pro ala cys cys asp arg val gln thr pro)
    cys arg arg ala val arg arg gly gly arg leu arg ala ala thr gly ser arg arg arg
1141/381                                1171/391
gtg aca ccg gcc gcc gat atc agt gcc cac cag ctg gtg cgg gcc gtg gta gac acc ggc
val thr pro ala ala asp ile ser ala his gln leu val arg ala val val asp thr gly
  OPA(his arg pro pro ile ser val pro thr ser trp cys gly pro trp)AMB(thr pro ala
    asp thr gly arg arg tyr gln cys pro pro ala gly ala gly arg gly arg)his arg arg
1201/401      SEQ ID NO: 709             1231/411     SEQ ID NO: 710
gcc gcg cac gtg atg gtg ctg ccc aat ggc tat gtg gcc gcc gaa gaa ctg gtg gcc ggg
ala ala his val met val leu pro asn gly tyr val ala ala glu glu leu val ala gly
  pro arg thr)OPA(trp cys cys pro met ala met trp pro pro lys asn trp trp pro gly
    arg ala arg asp)gly ala ala gln trp leu cys gly arg arg arg thr gly gly arg val
1261/421       SEQ ID NO: 711            1291/431
tgt acc gcg gcg atc ggc tgg ggc gtc gac gtg gta ccc gtg ccg acc gga tcg atg gtg
cys thr ala ala ile gly trp gly val asp val val pro val pro thr gly ser met val
  val pro arg arg ser ala gly ala ser thr trp tyr pro cys arg pro asp arg trp cys
    tyr arg gly asp arg leu gly arg arg arg gly thr arg ala asp arg ile asp gly ala
1321/441                                1351/451
cag ggg ttg gcc gcg ctg gcc gtg cat gac gcg gcc cgc cag gcc gtc gac gac ggc tac
gln gly leu ala ala leu ala val his asp ala ala arg gln ala val asp asp gly tyr
  arg gly trp pro arg trp pro cys met thr arg pro ala arg pro ser thr thr ala thr
    gly val gly arg ala gly arg ala)OPA(arg gly pro pro gly arg arg arg arg leu gln
1381/461       SEQ ID NO: 727            1411/471
agc atg gcc cgt gcc gcc ggt gct tcc cgg cac gga tcg gtg cgc att gcc acc caa aag
ser met ala arg ala ala gly ala ser arg his gly ser val arg ile ala thr gln lys
  ala trp pro val pro pro val leu pro gly thr asp arg cys ala leu pro pro lys arg
    his gly pro cys arg arg cys phe pro ala arg ile gly ala his cys his pro lys gly
```

SEQ ID NOS:697-727 (continued 1)

FIG. 41G (continued 1)

```
1441/481                                      1471/491
gcg ctg acc tgg gcc ggt acc tgc aag ccg ggc gac ggt ctg ggt atc gcg ggc gac gag
ala leu thr trp ala gly thr cys lys pro gly asp gly leu gly ile ala gly asp glu
 arg)OPA(pro gly pro val pro ala ser arg ala thr val trp val ser arg ala thr arg
  ala asp\leu gly arg tyr leu gln ala gly arg arg ser gly tyr arg gly arg arg gly
1501/501   \SEQ ID NO: 712                    1531/511
gtg ctg atc gtc gcc gac gat gtc gcc gcg gcg gcc atc ggt ctg gtc gac ctg ttg ttg
val leu ile val ala asp asp val ala ala ala ala ile gly leu val asp leu leu leu
 cys)OPA(ser ser pro thr met ser pro arg arg pro ser val trp ser thr cys cys trp
  ala asp\arg arg arg arg cys arg arg gly gly his arg ser gly arg pro val val gly
1561/521   \SEQ ID NO: 713                    1591/531
gca tcg gga ggc gat ctg gtg acg gtg cta att ggc gcc ggc gta acc gaa gac gtg gct
ala ser gly gly asp leu val thr val leu ile gly ala gly val thr glu asp val ala
 his arg glu ala ile trp)OPA arg cys OCH(leu ala pro ala)OCH(pro lys thr trp leu
  ile gly arg arg ser gly asp gly ala/asn trp arg arg arg asn\arg arg arg gly cys
1621/541            SEQ ID NO: 714-----       1651/551              \SEQ ID NO: 715
gtc gtc ctg gaa cgg cat gtg cac gac cac cat cca ggc acc gag ctg gtc tcc tac cgc
val val leu glu arg his val his asp his pro gly thr glu leu val ser tyr arg
 ser ser trp asn gly met cys thr thr thr ile gln ala pro ser trp ser pro thr ala
  arg pro gly thr ala cys ala arg pro pro ser arg his arg ala gly leu leu pro his
1681/561                                      1711/571
acc gga cac cgc ggc gac gcg ctg ctg atc ggg gtc gag tag
thr gly his arg gly asp ala leu leu ile gly val glu)AMB
 pro asp thr ala ala thr arg cys)OPA(ser gly ser)ser
  arg thr pro arg arg arg ala ala asp)arg gly arg val
              SEQ ID NO: 716-/
```

SEQ ID NOS:697-727 (continued 2)

FIG. 41G (continued 2)

```
1/1 _____-SEQ ID NO: 728                     31/11
GCC GGT AAC GCC GCG TCC CAG TGC TAT CCG TCC GCC GGA CCG CCC GAA ACA TCA GCG GCG
(ala gly asn ala ala ser gln cys tyr pro ser ala gly pro pro glu thr ser ala ala
61/21 ---SEQ ID NO: 729                       91/31
GGC GCC CCG GTC GGC CGC GGC CGG GCT CGA CCC GCT CCA CCT GGC CAT CAG CGA CCA GGT
gly ala pro val gly arg gly arg ala arg pro ala pro pro gly his gln arg pro gly
121/41                                        151/51
TAT CGA GGT GGA AGC GGA CGG TGT TGG GAT GCA CGC CCA ACT TGC CGG CGA TCG CGG CGA
tyr arg gly gly ser gly arg cys trp asp ala arg pro thr cys arg arg ser arg arg
181/61                                        211/71
TGC TCA TCG GAA CCC GCG ACG CAC ACA ATG CCC GCA GCA CCG CAC GAC GGC GCC CCA CCG
cys ser ser glu pro ala thr his thr met pro ala ala pro his asp gly ala pro pro
241/81                                        271/91
GCT CTT GCA GTG ACC TGA TGA TGA CAC TCA CCC CCA TAA GGC TCG TCG GCT GCG CCT GAG
ala leu ala val thr)OPA OPA OPA(his ser pro pro)OCH(gly ser ser ala ala pro glu
301/101            SEQ ID NO:730 --/          331/111           \-SEQ ID NO: 731
CAA TGC AGT AAG TTT ACA CAA ACG GAC TTG TAA AAA CCT GCG GAG GTG GGG TCT ATG GCC
gln cys ser lys phe thr gln thr asp leu)OCH(lys pro ala glu val gly ser met ala
361/121                                       391/131 SEQ ID NO: 732
AAC AAA CGT GGC AAT GCC GGG CAG CCT CTG CCC TTG TCG GAT C
asn lys arg gly asn ala gly gln pro leu pro leu ser asp)
```

SEQ ID NOS:728-732

FIG. 42A

1

1/1 SEQ ID NO: 733   31/11
CCG GTA ACG CCG CGT CCC AGT GCT ATC CGT CCG CCG GAC CGC CCG AAA CAT CAG CGG CGG
(pro val thr pro arg pro ser ala ile arg pro pro asp arg pro lys his gln arg arg
61/21   SEQ ID NO: 734   91/31
GCG CCC CGG TCG GCC GCG GCC GGG CTC GAC CCG CTC CAC CTG GCC ATC AGC GAC CAG GTT
ala pro arg ser ala ala ala gly leu asp pro leu his leu ala ile ser asp gln val
121/41   151/51
ATC GAG GTG GAA GCG GAC GGT GTT GGG ATG CAC GCC CAA CTT GCC GGC GAT CGC GGC GAT
ile glu val glu ala asp gly val gly met his ala gln leu ala gly asp arg gly asp
181/61   211/71
GCT CAT CGG AAC CCG CGA CGC ACA CAA TGC CCG CAG CAC CGC ACG ACG GCG CCC CAC CGG
ala his arg asn pro arg arg thr gln cys pro gln his arg thr thr ala pro his arg
241/81   271/91
CTC TTG CAG TGA CCT GAT GAT GAC ACT CAC CCC CAT AAG GCT CGT CGG CTG CGC CTG AGC
leu leu gln)OPA(pro asp asp asp thr his pro his lys ala arg arg leu arg leu ser
301/101   SEQ ID NO: 735   331/111
AAT GCA GTA AGT TTA CAC AAA CGG ACT TGT AAA AAC CTG CGG AGG TGG GGT CTA TGG CCA
asn ala val ser leu his lys arg thr cys lys asn leu arg arg trp gly leu trp pro
361/121   391/131
ACA AAC GTG GCA ATG CCG GGC AGC CTC TGC CCT TGT CGG ATC
thr asn val ala met pro gly ser leu cys pro cys arg ile)

SEQ ID NOS:733-735

FIG. 42B

1/1 SEQ ID NO: 736   31/11
CGG TAA CGC CGC GTC CCA GTG CTA TCC GTC CGC CGG ACC GCC CGA AAC ATC AGC GGC GGG
arg OCH(arg arg val pro val leu ser val arg arg thr ala arg asn ile ser gly gly
61/21   SEQ ID NO: 737   91/31
CGC CCC GGT CGG CCG CGG CCG GGC TCG ACC CGC TCC ACC TGG CCA TCA GCG ACC AGG TTA
arg pro gly arg pro arg pro gly ser thr arg ser thr trp pro ser ala thr arg leu
121/41   151/51
TCG AGG TGG AAG CGG ACG GTG TTG GGA TGC ACG CCC AAC TTG CCG GCG ATC GCG GCG ATG
ser arg trp lys arg thr val leu gly cys thr pro asn leu pro ala ile ala ala met
181/61   211/71
CTC ATC GGA ACC CGC GAC GCA CAC AAT GCC CGC AGC ACC GCA CGA CGG CGC CCC ACC GGC
leu ile gly thr arg asp ala his asn ala arg ser thr ala arg arg arg pro thr gly
241/81   271/91
TCT TGC AGT GAC CTG ATG ATG ACA CTC ACC CCC ATA AGG CTC GTC GGC TGC GCC TGA GCA
ser cys ser asp leu met met thr leu thr pro ile arg leu val gly cys ala)OPA ala
301/101   331/111
ATG CAG TAA GTT TAC ACA AAC GGA CTT GTA AAA ACC TGC GGA GGT GGG GTC TAT GGC CAA
met gln OCH(val tyr thr asn gly leu val lys thr cys gly gly gly val tyr gly gln
361/121   SEQ ID NO: 738   391/131
CAA ACG TGG CAA TGC CGG GCA GCC TCT GCC CTT GTC GGA TC
gln thr trp gln cys arg ala ala ser ala leu val gly)

SEQ ID NOS:736-738

FIG. 42C

Coding sequence Rv2622 predicted by Cole et al., 1998 (Nature 393:537-544) and containing seq42A:

1/1 SEQ ID NO: 739   31/11
atg gcc aac aaa cgt ggc aat gcc ggg cag cct ctg ccc ttg tcg gat cga gac gac gac
(Met ala asn lys arg gly asn ala gly gln pro leu pro leu ser asp arg asp asp asp
61/21 SEQ ID NO: 740   91/31
cac atg cag ggg cac tgg ctg ctg gcc cgg ctg ggc aag cgg gtg ctg cgt ccc ggc ggc
his met gln gly his trp leu leu ala arg leu gly lys arg val leu arg pro gly gly
121/41   151/51
gtc gaa ctc acc cgg aca ctg ctg gcc cgc gcc gag gtg acc gac gcc gac gtg ctc gag
val glu leu thr arg thr leu leu ala arg ala glu val thr asp ala asp val leu glu
181/61   211/71
ctg gca ccg ggc ctg ggc cgc acc gca gcc gaa atc ttg gcc cgc aac ccg cgg tcg tac
leu ala pro gly leu gly arg thr ala ala glu ile leu ala arg asn pro arg ser tyr
241/81   271/91
gtg ggg gcg gag agc gat ccc aac gcg gcc aac ctg gtc cga cac gtt ctc gcc ggc cgc
val gly ala glu ser asp pro asn ala ala asn leu val arg his val leu ala gly arg
301/101   331/111
ggc gac gtc cgg gtc acc gac gcg gcc gat acc gga tta tcc gac gcc agc gcc gat gtc
gly asp val arg val thr asp ala ala asp thr gly leu ser asp ala ser ala asp val
361/121   391/131
gtc atc ggc gag gcg atg ctg acc atg caa ggc aac gcg gct aaa cac acg atc gtc gcc
val ile gly glu ala met leu thr met gln gly asn ala ala lys his thr ile val ala
421/141   451/151
gag gcg gcg cgg gtg ctg agg ccg ggt ggc cgc tac gcg att cac gaa cta gcg ctg gtg
glu ala ala arg val leu arg pro gly gly arg tyr ala ile his glu leu ala leu val
481/161   511/171
ccg gac gac gtc gca gag cag gtc cgc acc gac ctg cgg cag tcg ctg gcc cgc gcg ctc
pro asp asp val ala glu gln val arg thr asp leu arg gln ser leu ala arg ala leu
541/181   571/191
aag gtc aat gcg cgt ccg ctg acc gtt gcg gaa tgg tcg cac ctc tta gcg ggc cat gga
lys val asn ala arg pro leu thr val ala glu trp ser his leu leu ala gly his gly
601/201   631/211
ctg gtc gtc gaa cac gtt gtc acc gct tcc atg gcg ttg tta caa ccg cga cgg gtg atc
leu val val glu his val val thr ala ser met ala leu leu gln pro arg arg val ile
661/221   691/231
gct gac gaa ggc ctc ctg ggt gcg ctg cgg ttc gcc gga aac ctg ctc atc cat cgt gcc
ala asp glu gly leu leu gly ala leu arg phe ala gly asn leu leu ile his arg ala
721/241   751/251
gcg cgt cgg cga gtc ctg ttg atg cgc cac aca ttc cgc agg cat cgt gaa cgc ttg aca
ala arg arg arg val leu leu met arg his thr phe arg arg his arg glu arg leu thr
781/261   811/271
gcc gtc gcc att gtc gcg cac aaa ccg cac gtc gat tcg tga
ala val ala ile val ala his lys pro his val asp ser)OPA

SEQ ID NOS:739-740

FIG. 42D

ORF according to Cole et al., 1998 (Nature 393:537-544) and containing Rv2622

```
1/1   SEQ ID NO: 741                        31/11
taa aaa cct gcg gag gtg ggg tct atg gcc    aac aaa cgt ggc aat gcc ggg cag cct ctg
OCH(lys pro ala glu val gly ser met ala    asn lys arg gly asn ala gly gln pro leu
61/21   SEQ ID NO: 742                      91/31
ccc ttg tcg gat cga gac gac gac cac atg    cag ggg cac tgg ctg ctg gcc cgg ctg ggc
pro leu ser asp arg asp asp asp his met    gln gly his trp leu leu ala arg leu gly
121/41                                      151/51
aag cgg gtg ctg cgt ccc ggc ggc gtc gaa    ctc acc cgg aca ctg ctg gcc cgc gcc gag
lys arg val leu arg pro gly gly val glu    leu thr arg thr leu leu ala arg ala glu
181/61                                      211/71
gtg acc gac gcc gac gtg ctc gag ctg gca    ccg ggc ctg ggc cgc acc gca gcc gaa atc
val thr asp ala asp val leu glu leu ala    pro gly leu gly arg thr ala ala glu ile
241/81                                      271/91
ttg gcc cgc aac ccg cgg tcg tac gtg ggg    gcg gag agc gat ccc aac gcg gcc aac ctg
leu ala arg asn pro arg ser tyr val gly    ala glu ser asp pro asn ala ala asn leu
301/101                                     331/111
gtc cga cac gtt ctc gcc ggc cgc ggc gac    gtc cgg gtc acc gac gcg gcc gat acc gga
val arg his val leu ala gly arg gly asp    val arg val thr asp ala ala asp thr gly
361/121                                     391/131
tta tcc gac gcc agc gcc gat gtc gtc atc    ggc gag gcg atg ctg acc atg caa ggc aac
leu ser asp ala ser ala asp val val ile    gly glu ala met leu thr met gln gly asn
421/141                                     451/151
gcg gct aaa cac acg atc gtc gcc gag gcg    gcg cgg gtg ctg agg ccg ggt ggc cgc tac
ala ala lys his thr ile val ala glu ala    ala arg val leu arg pro gly gly arg tyr
481/161                                     511/171
gcg att cac gaa cta gcg ctg gtg ccg gac    gac gtc gca gag cag gtc cgc acc gac ctg
ala ile his glu leu ala leu val pro asp    asp val ala glu gln val arg thr asp leu
541/181                                     571/191
cgg cag tcg ctg gcc cgc gcg ctc aag gtc    aat gcg cgt ccg ctg acc gtt gcg gaa tgg
arg gln ser leu ala arg ala leu lys val    asn ala arg pro leu thr val ala glu trp
601/201                                     631/211
tcg cac ctc tta gcg ggc cat gga ctg gtc    gtc gaa cac gtt gtc acc gct tcc atg gcg
ser his leu leu ala gly his gly leu val    val glu his val val thr ala ser met ala
661/221                                     691/231
ttg tta caa ccg cga cgg gtg atc gct gac    gaa ggc ctc ctg ggt gcg ctg cgg ttc gcc
leu leu gln pro arg arg val ile ala asp    glu gly leu leu gly ala leu arg phe ala
721/241                                     751/251
gga aac ctg ctc atc cat cgt gcc gcg cgt    cgg cga gtc ctg ttg atg cgc cac aca ttc
gly asn leu leu ile his arg ala ala arg    arg arg val leu leu met arg his thr phe
781/261                                     811/271
cgc agg cat cgt gaa cgc ttg aca gcc gtc    gcc att gtc gcg cac aaa ccg cac gtc gat
arg arg his arg glu arg leu thr ala val    ala ile val ala his lys pro his val asp
841/281
tcg tga
ser)OPA
```

SEQ ID NOS:741-742

FIG. 42E

1/1 SEQ ID NO: 743   31/11
atc gcg cgt gac atc gat gac cag ggt cgg ctg tgt ctg gac gtc ggc ggt cga acg gta
(ile ala arg asp ile asp asp gln gly arg leu cys leu asp val gly gly arg thr val
61/21 SEQ ID NO: 744   91/31
gtt gtt tca gcg ggc gac gtg gtg cat ttg cgt taa ctc gcg cgg agc tgg cgt ccc caa
val val ser ala gly asp val val his leu arg)OCH(leu ala arg ser trp arg pro gln
121/41   151/51 SEQ ID NO: 745
aag att aag gtc gcg ggc atg agc tat ccg gag aat gtc ctg gcc gct ggc gag cag gtc
lys ile lys val ala gly met ser tyr pro glu asn val leu ala ala gly glu gln val
181/61   211/71
gtt ctg cac cgc cat ccg cac tgg aat cgc tta atc tgg ccc gtc gtg gtg ctg gtc ttg
val leu his arg his pro his trp asn arg leu ile trp pro val val val leu val leu
241/81   271/91
ctg acc ggg ttg gcg gcg ttc ggg tcc gga ttc gtc aac tcg aca cct tgg cag cag atc
leu thr gly leu ala ala phe gly ser gly phe val asn ser thr pro trp gln gln ile)

SEQ ID NOS:743-745

FIG. 43A

1/1 SEQ ID NO: 746   31/11
tcg cgc gtg aca tcg atg acc agg gtc ggc tgt gtc tgg acg tcg gcg gtc gaa cgg tag
(ser arg val thr ser met thr arg val gly cys val trp thr ser ala val glu arg)AMB
61/21 SEQ ID NO: 747   91/31
ttg ttt cag cgg gcg acg tgg tgc att tgc gtt aac tcg cgc gga gct ggc gtc ccc aaa
(leu phe gln arg ala thr trp cys ile cys val asn ser arg gly ala gly val pro lys
121/41 SEQ ID NO: 748   151/51
aga tta agg tcg cgg gca tga gct atc cgg aga atg tcc tgg ccg ctg gcg agc agg tcg
arg leu arg ser arg ala)OPA(ala ile arg arg met ser trp pro leu ala ser arg ser
181/61   SEQ ID NO: 749   211/71
ttc tgc acc gcc atc cgc act gga atc gct taa tct ggc ccg tcg tgg tgc tgg tct tgc
phe cys thr ala ile arg thr gly ile ala)OCH(ser gly pro ser trp cys trp ser cys)
241/81   271/91 SEQ ID NO: 750
tga ccg ggt tgg cgg cgt tcg ggt ccg gat tcg tca act cga cac ctt ggc agc aga tc
OPA(pro gly trp arg arg ser gly pro asp ser ser thr arg his leu gly ser arg)
       SEQ ID NO: 751

SEQ ID NOS:746-751

FIG. 43B

1/1 SEQ ID NO: 752   31/11
cgc gcg tga cat cga tga cca ggg tcg gct gtg tct gga cgt cgg cgg tcg aac ggt agt
arg ala OPA his arg OPA(pro gly ser ala val ser gly arg arg arg ser asn gly ser
61/21 SEQ ID NO: 753   91/31
tgt ttc agc ggg cga cgt ggt gca ttt gcg tta act cgc gcg gag ctg gcg tcc cca aaa
cys phe ser gly arg arg gly ala phe ala leu thr arg ala glu leu ala ser pro lys
121/41   151/51
gat taa ggt cgc ggg cat gag cta tcc gga gaa tgt cct ggc cgc tgg cga gca ggt cgt
asp)OCH(gly arg gly his glu leu ser gly glu cys pro gly arg trp arg ala gly arg
181/61   SEQ ID NO: 754   211/71
tct gca ccg cca tcc gca ctg gaa tcg ctt aat ctg gcc cgt cgt ggt gct ggt ctt gct
ser ala pro pro ser ala leu glu ser leu asn leu ala arg arg gly ala gly leu ala
241/81   271/91
gac cgg gtt ggc ggc gtt cgg gtc cgg att cgt caa ctc gac acc ttg gca gca gat c
asp arg val gly gly val arg val arg ile arg gln leu asp thr leu ala ala asp)

SEQ ID NOS:752-754

FIG. 43C

Coding sequence Rv3278c predicted by Cole et al., 1998 (Nature 393:537-544)
and containing seq43A:

1/1 SEQ ID NO: 755   31/11
atg agc tat ccg gag aat gtc ctg gcc gct ggc gag cag gtc gtt ctg cac cgc cat ccg
(Met ser tyr pro glu asn val leu ala ala gly glu gln val val leu his arg his pro
61/21 SEQ ID NO: 756   91/31
cac tgg aat cgc tta atc tgg ccc gtc gtg gtg ctg gtc ttg ctg acc ggg ttg gcg gcg
his trp asn arg leu ile trp pro val val val leu val leu leu thr gly leu ala ala
121/41   151/51
ttc ggg tcc gga ttc gtc aac tcg aca cct tgg cag cag atc gct aag aac gtg att cac
phe gly ser gly phe val asn ser thr pro trp gln gln ile ala lys asn val ile his
181/61   211/71
gcg gtc atc tgg ggg atc tgg ttg gtg atc gtc ggc tgg ctc acg ctg tgg cca ttc ctg
ala val ile trp gly ile trp leu val ile val gly trp leu thr leu trp pro phe leu
241/81   271/91
agc tgg ctg acc aca cat ttc gtg gtg acc aac cgg cgg gtg atg ttc cgg cat ggt gtg
ser trp leu thr thr his phe val val thr asn arg arg val met phe arg his gly val
301/101   331/111
ctg acc cgc agc ggg atc gac ata ccg cta gca cgg atc aac agc gtg gag ttc cgg gac
leu thr arg ser gly ile asp ile pro leu ala arg ile asn ser val glu phe arg asp
361/121   391/131
cgg atc ttc gag cgg att ttt cgc acc ggg acg ttg att atc gag tcc gcg tca caa gat
arg ile phe glu arg ile phe arg thr gly thr leu ile ile glu ser ala ser gln asp
421/141   451/151
ccg ctc gag ttc tac aac att ccg cgc ctg cgg gag gtg cat gcg ttg ctg tat cac gag
pro leu glu phe tyr asn ile pro arg leu arg glu val his ala leu leu tyr his glu
481/161   511/171
gtt ttc gac acc ctg ggc tcc gac gag tcg ccc agc tga
val phe asp thr leu gly ser asp glu ser pro ser)OPA

SEQ ID NOS:755-756

FIG. 43D

ORF according to Cole et al., 1998 (Nature 393:537-544) and containing Rv3278c

1/1 SEQ ID NO: 757   31/11
taa ctc gcg cgg agc tgg cgt ccc caa aag att aag gtc gcg ggc atg agc tat ccg gag
OCH(leu ala arg ser trp arg pro gln lys ile lys val ala gly met ser tyr pro glu
61/21 SEQ ID NO: 758   91/31
aat gtc ctg gcc gct ggc gag cag gtc gtt ctg cac cgc cat ccg cac tgg aat cgc tta
asn val leu ala ala gly glu gln val val leu his arg his pro his trp asn arg leu
121/41   151/51
atc tgg ccc gtc gtg gtg ctg gtc ttg ctg acc ggg ttg gcg gcg ttc ggg tcc gga ttc
ile trp pro val val val leu val leu leu thr gly leu ala ala phe gly ser gly phe
181/61   211/71
gtc aac tcg aca cct tgg cag cag atc gct aag aac gtg att cac gcg gtc atc tgg ggg
val asn ser thr pro trp gln gln ile ala lys asn val ile his ala val ile trp gly
241/81   271/91
atc tgg ttg gtg atc gtc ggc tgg ctc acg ctg tgg cca ttc ctg agc tgg ctg acc aca
ile trp leu val ile val gly trp leu thr leu trp pro phe leu ser trp leu thr thr
301/101   331/111
cat ttc gtg gtg acc aac cgg cgg gtg atg ttc cgg cat ggt gtg ctg acc cgc agc ggg
his phe val val thr asn arg arg val met phe arg his gly val leu thr arg ser gly
361/121   391/131
atc gac ata ccg cta gca cgg atc aac agc gtg gag ttc cgg gac cgg atc ttc gag cgg
ile asp ile pro leu ala arg ile asn ser val glu phe arg asp arg ile phe glu arg
421/141   451/151
att ttt cgc acc ggg acg ttg att atc gag tcc gcg tca caa gat ccg ctc gag ttc tac
ile phe arg thr gly thr leu ile ile glu ser ala ser gln asp pro leu glu phe tyr
481/161   511/171
aac att ccg cgc ctg cgg gag gtg cat gcg ttg ctg tat cac gag gtt ttc gac acc ctg
asn ile pro arg leu arg glu val his ala leu leu tyr his glu val phe asp thr leu
541/181
ggc tcc gac gag tcg ccc agc tga
gly ser asp glu ser pro ser)OPA

SEQ ID NOS:757-758

FIG. 43E

1/1 SEQ ID NO: 759   31/11
gcc aag atg gat gtc tac caa cgc acc gcc gcc ggc tgg cag ccg ctc aag acc ggt atc
(ala lys met asp val tyr gln arg thr ala ala gly trp gln pro leu lys thr gly ile
61/21 SEQ ID NO: 760   91/31
acc acc cat atc ggt tcg gcg ggc atg gcg ccg gaa gcc aag agc gga tat ccg gcc act
thr thr his ile gly ser ala gly met ala pro glu ala lys ser gly tyr pro ala thr
121/41   151/51
ccg atg ggg gtt tac agc ctg gac tcc gct ttt ggc acc gcg ccg aat ccc ggt ggc ggg
pro met gly val tyr ser leu asp ser ala phe gly thr ala pro asn pro gly gly gly
181/61   211/71
ttg ccg tat acc caa gtc gga ccc aat cac tgg tgg agt ggc gac gac aat agc ccc acc
leu pro tyr thr gln val gly pro asn his trp trp ser gly asp asp asn ser pro thr
241/81   271/91
ttt aac tcc atg cag gtc tgt cag aag tcc cag tgc ccg ttc agc acg gcc gac agc gag
phe asn ser met gln val cys gln lys ser gln cys pro phe ser thr ala asp ser glu
301/101   331/111
aac ctg caa atc ccg cag tac aag cat tcg gtc gtg atg ggc gtc aac aag gcc aag gtc
asn leu gln ile pro gln tyr lys his ser val val met gly val asn lys ala lys val
361/121   391/131
cca ggc aaa ggc tcc gcg ttc ttc ttt cac acc acc gac ggc ggg ccc acc gcg ggt tgt
pro gly lys gly ser ala phe phe phe his thr thr asp gly gly pro thr ala gly cys
421/141
gtg gcg atc
val ala ile)

SEQ ID NOS:759-760

FIG. 44A

1/1 SEQ ID NO: 761   31/11
cca aga tgg atg tct acc aac gca ccg ccg ccg gct ggc agc cgc tca aga ccg gta tca
(pro arg trp met ser thr asn ala pro pro pro ala gly ser arg ser arg pro val ser
61/21 SEQ ID NO: 762   91/31
cca ccc ata tcg gtt cgg cgg gca tgg cgc cgg aag cca aga gcg gat atc cgg cca ctc
pro pro ile ser val arg arg ala trp arg arg lys pro arg ala asp ile arg pro leu
121/41   151/51
cga tgg ggg ttt aca gcc tgg act ccg ctt ttg gca ccg cgc cga atc ccg gtg gcg ggt
arg trp gly phe thr ala trp thr pro leu leu ala pro arg arg ile pro val ala gly
181/61   211/71
tgc cgt ata ccc aag tcg gac cca atc act ggt gga gtg gcg acg aca ata gcc cca cct
cys arg ile pro lys ser asp pro ile thr gly gly val ala thr thr ile ala pro pro
241/81   271/91
tta act cca tgc agg tct gtc aga agt ccc agt gcc cgt tca gca cgg ccg aca gcg aga
leu thr pro cys arg ser val arg ser pro ser ala arg ser ala arg pro thr ala arg
301/101   331/111
acc tgc aaa tcc cgc agt aca agc att cgg tcg tga tgg gcg tca aca agg cca agg tcc
thr cys lys ser arg ser thr ser ile arg ser)OPA(trp ala ser thr arg pro arg ser
361/121   391/131 SEQ ID NO: 763
cag gca aag gct ccg cgt tct tct ttc aca cca ccg acg gcg ggc cca ccg cgg gtt gtg
gln ala lys ala pro arg ser ser phe thr pro pro thr ala gly pro pro arg val val
421/141
tgg cga tc
trp arg)

SEQ ID NOS:761-763

FIG. 44B

1/1 SEQ ID NO: 764   31/11
caa gat gga tgt cta cca acg cac cgc cgc cgg ctg gca gcc gct caa gac cgg tat cac
(gln asp gly cys leu pro thr his arg arg arg leu ala ala ala gln asp arg tyr his
61/21 SEQ ID NO: 765   91/31
cac cca tat cgg ttc ggc ggg cat ggc gcc gga agc caa gag cgg ata tcc ggc cac tcc
his pro tyr arg phe gly gly his gly ala gly ser gln glu arg ile ser gly his ser
121/41   151/51
gat ggg ggt tta cag cct gga ctc cgc ttt tgg cac cgc gcc gaa tcc cgg tgg cgg gtt
asp gly gly leu gln pro gly leu arg phe trp his arg ala glu ser arg trp arg val
181/61   211/71
gcc gta tac cca agt cgg acc caa tca ctg gtg gag tgg cga cga caa tag ccc cac ctt
ala val tyr pro ser arg thr gln ser leu val glu trp arg arg gln)AMB pro his leu
241/81   271/91
taa ctc cat gca ggt ctg tca gaa gtc cca gtg ccc gtt cag cac ggc cga cag cga gaa
OCH(leu his ala gly leu ser glu val pro val pro val gln his gly arg gln arg glu
301/101 SEQ ID NO: 766   331/111
cct gca aat ccc gca gta caa gca ttc ggt cgt gat ggg cgt caa caa ggc caa ggt ccc
pro ala asn pro ala val gln ala phe gly arg asp gly arg gln gln gly gln gly pro
361/121   391/131
agg caa agg ctc cgc gtt ctt ctt tca cac cac cga cgg cgg gcc cac cgc ggg ttg tgt
arg gln arg leu arg val leu leu ser his his arg arg arg ala his arg gly leu cys
421/141
ggc gat c
gly asp)

SEQ ID NOS:764-766

FIG. 44C

Coding sequence Rv0309 predicted by Cole et al., 1998 (Nature 393:537-544)
and containing Seq44A:

1/1 ___SEQ ID NO: 767   31/11
atg agc cga ctc cta gct ttg ctg tgc gct gcg gta tgc acg ggc tgc gtt gct gtg gtt
(Met ser arg leu leu ala leu leu cys ala ala val cys thr gly cys val ala val val
61/21 ~SEQ ID NO: 768   91/31
ctc gcg cca gtg agc ctg gcc gtc gtc aac ccg tgg ttc gcg aac tcg gtc ggc aat gcc
leu ala pro val ser leu ala val val asn pro trp phe ala asn ser val gly asn ala
121/41   151/51
act cag gtg gtt tcg gtg gtg gga acc ggc ggt tcg acg gcc aag atg gat gtc tac caa
thr gln val val ser val val gly thr gly gly ser thr ala lys met asp val tyr gln
181/61   211/71
cgc acc gcc gcc ggc tgg cag ccg ctc aag acc ggt atc acc acc cat atc ggt tcg gcg
arg thr ala ala gly trp gln pro leu lys thr gly ile thr thr his ile gly ser ala
241/81   271/91
ggc atg gcg ccg gaa gcc aag agc gga tat ccg gcc act ccg atg ggg gtt tac agc ctg
gly met ala pro glu ala lys ser gly tyr pro ala thr pro met gly val tyr ser leu
301/101   331/111
gac tcc gct ttt ggc acc gcg ccg aat ccc ggt ggc ggg ttg ccg tat acc caa gtc gga
asp ser ala phe gly thr ala pro asn pro gly gly gly leu pro tyr thr gln val gly
361/121   391/131
ccc aat cac tgg tgg agt ggc gac gac aat agc ccc acc ttt aac tcc atg cag gtc tgt
pro asn his trp trp ser gly asp asp asn ser pro thr phe asn ser met gln val cys
421/141   451/151
cag aag tcc cag tgc ccg ttc agc acg gcc gac agc gag aac ctg caa atc ccg cag tac
gln lys ser gln cys pro phe ser thr ala asp ser glu asn leu gln ile pro gln tyr
481/161   511/171
aag cat tcg gtc gtg atg ggc gtc aac aag gcc aag gtc cca ggc aaa ggc tcc gcg ttc
lys his ser val val met gly val asn lys ala lys val pro gly lys gly ser ala phe
541/181   571/191
ttc ttt cac acc acc gac ggc ggg ccc acc gcg ggt tgt gtg gcg atc gac gat gcc acg
phe phe his thr thr asp gly gly pro thr ala gly cys val ala ile asp asp ala thr
601/201   631/211~
ctg gtg cag atc atc cgt tgg ctg cgg cct ggt gcg gtg atc gcg atc gcc aag taa
leu val gln ile ile arg trp leu arg pro gly ala val ile ala ile ala lys)OCH

SEQ ID NOS:767-768

FIG. 44D

ORF according to Cole et al., 1998 (Nature 393:537-544) and containing Rv0309

```
1/1 SEQ ID NO: 769                          31/11
tga gcg atg agc cga ctc cta gct ttg ctg     tgc gct gcg gta tgc acg ggc tgc gtt gct
OPA(ala met ser arg leu leu ala leu leu     cys ala ala val cys thr gly cys val ala
61/21    SEQ ID NO: 770                      91/31
gtg gtt ctc gcg cca gtg agc ctg gcc gtc     gtc aac ccg tgg ttc gcg aac tcg gtc ggc
val val leu ala pro val ser leu ala val     val asn pro trp phe ala asn ser val gly
121/41                                       151/51
aat gcc act cag gtg gtt tcg gtg gtg gga     acc ggc ggt tcg acg gcc aag atg gat gtc
asn ala thr gln val val ser val val gly     thr gly gly ser thr ala lys met asp val
181/61                                       211/71
tac caa cgc acc gcc gcc ggc tgg cag ccg     ctc aag acc ggt atc acc acc cat atc ggt
tyr gln arg thr ala ala gly trp gln pro     leu lys thr gly ile thr thr his ile gly
241/81                                       271/91
tcg gcg ggc atg gcg ccg gaa gcc aag agc     gga tat ccg gcc act ccg atg ggg gtt tac
ser ala gly met ala pro glu ala lys ser     gly tyr pro ala thr pro met gly val tyr
301/101                                      331/111
agc ctg gac tcc gct ttt ggc acc gcg ccg     aat ccc ggt ggc ggg ttg ccg tat acc caa
ser leu asp ser ala phe gly thr ala pro     asn pro gly gly gly leu pro tyr thr gln
361/121                                      391/131
gtc gga ccc aat cac tgg tgg agt ggc gac     gac aat agc ccc acc ttt aac tcc atg cag
val gly pro asn his trp trp ser gly asp     asp asn ser pro thr phe asn ser met gln
421/141                                      451/151
gtc tgt cag aag tcc cag tgc ccg ttc agc     acg gcc gac agc gag aac ctg caa atc ccg
val cys gln lys ser gln cys pro phe ser     thr ala asp ser glu asn leu gln ile pro
481/161                                      511/171
cag tac aag cat tcg gtc gtg atg ggc gtc     aac aag gcc aag gtc cca ggc aaa ggc tcc
gln tyr lys his ser val val met gly val     asn lys ala lys val pro gly lys gly ser
541/181                                      571/191
gcg ttc ttc ttt cac acc acc gac ggc ggg     ccc acc gcg ggt tgt gtg gcg atc gac gat
ala phe phe phe his thr thr asp gly gly     pro thr ala gly cys val ala ile asp asp
601/201                                      631/211
gcc acg ctg gtg cag atc atc cgt tgg ctg     cgg cct ggt gcg gtg atc gcg atc gcc aag
ala thr leu val gln ile ile arg trp leu     arg pro gly ala val ile ala ile ala lys)
661/221
taa
OCH
```

SEQ ID NOS:769-770

FIG. 44E

Cloned fragment fused with phoA

1/1 SEQ ID NO: 771   31/11
gat ctc ccc gga cac cag gtc atc cgg cga gat ggt gat cga ggc tcg gac ccg cag gca
(asp leu pro gly his gln val ile arg arg asp gly asp arg gly ser asp pro gln ala
61/21 SEQ ID NO: 772   91/31
tcc ggt agc cag agg cac cag cat cag caa cat cgc gat ggc cag cat gcc gcg ccg tcg
ser gly ser gln arg his gln his gln gln his arg asp gly gln his ala ala pro ser
121/41   151/51
ggt cct tgc cac tcg cga tcc ttg gga tga cgg tgg ggc ata gct agc gcg cac cag gtc
gly pro cys his ser arg ser leu gly)OPA(arg trp gly ile ala ser ala his gln val
181/61   ID NO: 773   211/71
atc gtg cca gac cgg gca tgc cgc gtc ggc aag ctg tcg ggc gcg ggt tag agc ggt agc
ile val pro asp arg ala cys arg val gly lys leu ser gly ala gly)AMB(ser gly ser
241/81   271/91   SEQ ID NO: 774
gtg cga ccc agg atg gcg aat gct cgg ggg tca ccg gcg aag tgg tag ccg cgg atg atg
val arg pro arg met ala asn ala arg gly ser pro ala lys trp)AMB(pro arg met met
301/101   331/111   SEQ ID NO: 775
tcg gtg aag ccc aac cgg cgg tac aac cgc cac gcc cga ttg tcc tca ccg ttg gtc tcc
ser val lys pro asn arg arg tyr asn arg his ala arg leu ser ser pro leu val ser
361/121   391/131
ggt gtg gag agc agg acg ttg tcc tcg tcg cga ccg gct agc agt cgg cgg gcc aac gcc
gly val glu ser arg thr leu ser ser ser arg pro ala ser ser arg arg ala asn ala
421/141   451/151
tcc ccg agg cca cgg cct tga gcg cgg gga agg atg tgc aat tca gtc aac tcg aag tag
ser pro arg pro arg pro)OPA(ala arg gly arg met cys asn ser val asn ser lys)AMB
481/161   SEQ ID NO: 776   511/171
ctg gtc atc agt cgg gcg atc gct agg cgc gga aag ccg ctg cgt tgc aag ccc agt acc
(leu val ile ser arg ala ile ala arg arg gly lys pro leu arg cys lys pro ser thr
541/181 SEQ ID NO: 777   571/191
acc tgc tgt tgc cac cac tgg ccg ggc gcc ccg gga tag ccg tac gcc act ccg agc att
thr cys cys cys his his trp pro gly ala pro gly)AMB(pro tyr ala thr pro ser ile
601/201   631/211   SEQ ID NO: 778
ggc gcg ttg ctc agt tcg gcg gcc gac ggc agc gcc gtg gtg tcg gcg gcc tcg gcc tgt
gly ala leu leu ser ser ala ala asp gly ser ala val val ser ala ala ser ala cys
661/221   691/231
tcg gct gcc gtt acc tcg acg gcc gcg acc gcc tgc cag ccg cgc cgc cgg atg tgc tcc
ser ala ala val thr ser thr ala ala thr ala cys gln pro arg arg arg met cys ser
721/241   751/251
agc cac att ggg gcg cgc aaa gtc tcg gtg ccc ctg ggg tag cgc atc gcg tcg aca tac
ser his ile gly ala arg lys val ser val pro leu gly)AMB(arg ile ala ser thr tyr
781/261   811/271   SEQ ID NO: 779
acc gtc agg gca tca ccg agg cgg cgc tcc ata tcg ctg ggc ggc aga tcg atg agg aat
thr val arg ala ser pro arg arg arg ser ile ser leu gly gly arg ser met arg asn
841/281   871/291
atc gcc aac gcg cgg tgt cct cct cat gtg atg aac cga tgc gtg ctt gcg cac cag tat
ile ala asn ala arg cys pro pro his val met asn arg cys val leu ala his gln tyr
901/301   931/311
cgg aca agc cga tga ggc cgc ccg cgc tgg acg ggg ctt gta gcg tat ggc cgt ttc cgc
arg thr ser arg)OPA(gly arg pro arg trp thr gly leu val ala tyr gly arg phe arg
                 SEQ ID NO: 780

SEQ ID NOS:771-783

FIG. 45ZA

961/321
tca gct cgt cgc tgc ggc gcc gcc ggg ata
ser ala arg arg cys gly ala ala gly ile
1021/341   SEQ ID NO: 782
gat tga cct cgt atc atc tga gtt agt tgc
asp)OPA(pro arg ile ile)OPA(val ser cys
1081/361   SEQ ID NO: 781
att acg tga cag tct gtc ggc aag gag gga
ile thr)OPA(gln ser val gly lys glu gly
1141/381   SEQ ID NO: 783
atg ctt gac cag atc gag agc gct ctc tac
met leu asp gln ile glu ser ala leu tyr
1201/401
cgt ggc ggg ggc ttc cgc gca ccg acc gcg
arg gly gly gly phe arg ala pro thr ala
1261/421
atc atc ggt ctg ggg atg ttg gtt tcc ggc
ile ile gly leu gly met leu val ser gly
1321/441
ttc ccg ata ctc agc gtt ttc ggt ttt gtc
phe pro ile leu ser val phe gly phe val
1381/461
acc ggt cct cgg ttg tcc ggc agg atg gat
thr gly pro arg leu ser gly arg met asp
1441/481
cag cgt cgt acc aag ggg gcc ggg ggc tca
gln arg arg thr lys gly ala gly gly ser

991/331
gaa tcg ccc gcg aac cag tgg tac ggc gca
glu ser pro ala asn gln trp tyr gly ala
1051/351
ccg cgc aat ggg cat ccg cgt gtt atc ggt
pro arg asn gly his pro arg val ile gly
1111/371
cgc atg cca ctc tcc gat cat gag cag cgg
arg met pro leu ser asp his glu gln arg
1171/391
gcc gaa gat ccc aag ttc gca tcg agt gtc
ala glu asp pro lys phe ala ser ser val
1231/411
cgg cgg cgc ctg cag ggc gcg gcg ttg ttc
arg arg arg leu gln gly ala ala leu phe
1291/431
gtg gcg ttc aaa gag acc atg atc gga agt
val ala phe lys glu thr met ile gly ser
1351/451
gtg atg ttc ggt ggt gtg gtg tat gcc atc
val met phe gly gly val val tyr ala ile
1411/471
cgt ggc gga tcg gct gct ggg gct tcg cgc
arg gly gly ser ala ala gly ala ser arg
1471/491
ttc acc agc cgt atg gaa gat c
phe thr ser arg met glu asp)

SEQ ID NOS:771-783 (continued)

FIG. 45ZA (continued)

fragment seq45ZA shifted minus 1 for the reading frame

1/1 SEQ ID NO: 784   31/11
atc tcc ccg gac acc agg tca tcc ggc gag atg gtg atc gag gct cgg acc cgc agg cat
(ile ser pro asp thr arg ser ser gly glu met val ile glu ala arg thr arg arg his
61/21 SEQ ID NO: 785   91/31
ccg gta gcc aga ggc acc agc atc agc aac atc gcg atg gcc agc atg ccg cgc cgt cgg
pro val ala arg gly thr ser ile ser asn ile ala met ala ser met pro arg arg arg
121/41   151/51
gtc ctt gcc act cgc gat cct tgg gat gac ggt ggg gca tag cta gcg cgc acc agg tca
val leu ala thr arg asp pro trp asp asp gly gly ala)AMB(leu ala arg thr arg ser
181/61   211/71   SEQ ID NO: 786
tcg tgc cag acc ggg cat gcc gcg tcg gca agc tgt cgg gcg cgg gtt aga gcg gta gcg
ser cys gln thr gly his ala ala ser ala ser cys arg ala arg val arg ala val ala
241/81   271/91
tgc gac cca gga tgg cga atg ctc ggg ggt cac cgg cga agt ggt agc cgc gga tga tgt
cys asp pro gly trp arg met leu gly gly his arg arg ser gly ser arg gly)OPA cys
301/101   331/111
cgg tga agc cca acc ggc ggt aca acc gcc acg ccc gat tgt cct cac cgt tgg tct ccg
arg OPA(ser pro thr gly gly thr thr ala thr pro asp cys pro his arg trp ser pro
361/121 SEQ ID NO: 787   391/131
gtg tgg aga gca gga cgt tgt cct cgt cgc gac cgg cta gca gtc ggc ggg cca acg cct
val trp arg ala gly arg cys pro arg arg asp arg leu ala val gly gly pro thr pro
421/141   451/151
ccc cga ggc cac ggc ctt gag cgc ggg gaa gga tgt gca att cag tca act cga agt agc
pro arg gly his gly leu glu arg gly glu gly cys ala ile gln ser thr arg ser ser
481/161   511/171
tgg tca tca gtc ggg cga tcg cta ggc gcg gaa agc cgc tgc gtt gca agc cca gta cca
trp ser ser val gly arg ser leu gly ala glu ser arg cys val ala ser pro val pro
541/181   571/191
cct gct gtt gcc acc act ggc cgg gcg ccc cgg gat agc cgt acg cca ctc cga gca ttg
pro ala val ala thr thr gly arg ala pro arg asp ser arg thr pro leu arg ala leu
601/201   631/211
gcg cgt tgc tca gtt cgg cgg ccg acg gca gcg ccg tgg tgt cgg cgg cct cgg cct gtt
ala arg cys ser val arg arg pro thr ala ala pro trp cys arg arg pro arg pro val
661/221   691/231
cgg ctg ccg tta cct cga cgg ccg cga ccg cct gcc agc cgc gcc gcc gga tgt gct cca
arg leu pro leu pro arg arg pro arg pro pro ala ser arg ala ala gly cys ala pro
721/241   751/251
gcc aca ttg ggg cgc gca aag tct cgg tgc ccc tgg ggt agc gca tcg cgt cga cat aca
ala thr leu gly arg ala lys ser arg cys pro trp gly ser ala ser arg arg his thr
781/261   811/271
ccg tca ggg cat cac cga ggc ggc gct cca tat cgc tgg gcg gca gat cga tga gga ata
pro ser gly his his arg gly gly ala pro tyr arg trp ala ala asp arg)OPA(gly ile
841/281   871/291   SEQ ID NO: 788
tcg cca acg cgc ggt gtc ctc ctc atg tga tga acc gat gcg tgc ttg cgc acc agt atc
ser pro thr arg gly val leu leu met)OPA OPA(thr asp ala cys leu arg thr ser ile
901/301   931/311 SEQ ID NO: 789
gga caa gcc gat gag gcc gcc cgc gct gga cgg ggc ttg tag cgt atg gcc gtt tcc gct
gly gln ala asp glu ala ala arg ala gly arg gly leu)AMB(arg met ala val ser ala
961/321   991/331   SEQ ID NO: 790
cag ctc gtc gct gcg gcg ccg ccg gga tag aat cgc ccg cga acc agt ggt acg gcg cag
gln leu val ala ala ala pro pro gly)AMB(asn arg pro arg thr ser gly thr ala gln
                                          SEQ ID NO: 791

SEQ ID NOS:784-793

FIG. 45ZB

1021/341
att gac ctc gta tca tct gag tta gtt gcc
ile asp leu val ser ser glu leu val ala
1051/351
cgc gca atg ggc atc cgc gtg tta tcg gta
arg ala met gly ile arg val leu ser val

1081/361
tta cgt gac agt ctg tcg gca agg agg gac
leu arg asp ser leu ser ala arg arg asp
1111/371
gca tgc cac tct ccg atc atg agc agc gga
ala cys his ser pro ile met ser ser gly

1141/381
tgc ttg acc aga tcg aga gcg ctc tct acg
cys leu thr arg ser arg ala leu ser thr
1171/391
ccg aag atc cca agt tcg cat cga gtg tcc
pro lys ile pro ser ser his arg val ser

1201/401
gtg gcg ggg gct tcc gcg cac cga ccg cgc
val ala gly ala ser ala his arg pro arg
1231/411
ggc ggc gcc tgc agg gcg cgg cgt tgt tca
gly gly ala cys arg ala arg arg cys ser

1261/421
tca tcg gtc tgg gga tgt tgg ttt ccg gcg
ser ser val trp gly cys trp phe pro ala
1291/431
tgg cgt tca aag aga cca tga tcg gaa gtt
trp arg ser lys arg pro)OPA(ser glu val

1321/441
tcc cga tac tca gcg ttt tcg gtt ttg tcg
ser arg tyr ser ala phe ser val leu ser
1351/451   SEQ ID NO: 792
tga tgt tcg gtg gtg tgg tgt atg cca tca
)OPA(cys ser val val trp cys met pro ser

1381/461
ccg gtc ctc ggt tgt ccg gca gga tgg atc
pro val leu gly cys pro ala gly trp ile
1411/471   SEQ ID NO: 793
gtg gcg gat cgg ctg ctg ggg ctt cgc gcc
val ala asp arg leu leu gly leu arg ala

1441/481
agc gtc gta cca agg ggg ccg ggg gct cat
ser val val pro arg gly pro gly ala his
1471/491
tca cca gcc gta tgg aag atc
ser pro ala val trp lys ile)

SEQ ID NOS:784-793 (continued)

FIG. 45ZB (continued)

fragment seq45ZA shifted minus 2 for the reading frame

1/1 ⌐SEQ ID NO: 794   31/11
tct ccc cgg aca cca ggt cat ccg gcg aga   tgg tga tcg agg ctc gga ccc gca ggc atc
(ser pro arg thr pro gly his pro ala arg   trp)OPA(ser arg leu gly pro ala gly ile
61/21   ⌐SEQ ID NO: 795   91/31   ⌐SEQ ID NO: 796
cgg tag cca gag gca cca gca tca gca aca   tcg cga tgg cca gca tgc cgc gcc gtc ggg
arg)AMB(pro glu ala pro ala ser ala thr   ser arg trp pro ala cys arg ala val gly
121/41   ⌐SEQ ID NO: 797   151/51
tcc ttg cca ctc gcg atc ctt ggg atg acg   gtg ggg cat agc tag cgc gca cca ggt cat
ser leu pro leu ala ile leu gly met thr   val gly his ser)AMB(arg ala pro gly his
181/61   211/71   ⌐SEQ ID NO: 798
cgt gcc aga ccg ggc atg ccg cgt cgg caa   gct gtc ggg cgc ggg tta gag cgg tag cgt
arg ala arg pro gly met pro arg arg gln   ala val gly arg gly leu glu arg)AMB(arg
241/81   271/91   SEQ ID NO: 799 ⌐
gcg acc cag gat ggc gaa tgc tcg ggg gtc   acc ggc gaa gtg gta gcc gcg gat gat gtc
ala thr gln asp gly glu cys ser gly val   thr gly glu val val ala ala asp asp val
301/101   331/111
ggt gaa gcc caa ccg gcg gta caa ccg cca   cgc ccg att gtc ctc acc gtt ggt ctc cgg
gly glu ala gln pro ala val gln pro pro   arg pro ile val leu thr val gly leu arg
361/121   391/131
tgt gga gag cag gac gtt gtc ctc gtc gcg   acc ggc tag cag tcg gcg ggc caa cgc ctc
cys gly glu gln asp val val leu val ala   thr gly)AMB(gln ser ala gly gln arg leu
421/141   451/151   ⌐SEQ ID NO: 800
ccc gag gcc acg gcc ttg agc gcg ggg aag   gat gtg caa ttc agt caa ctc gaa gta gct
pro glu ala thr ala leu ser ala gly lys   asp val gln phe ser gln leu glu val ala
481/161   511/171
ggt cat cag tcg ggc gat cgc tag gcg cgg   aaa gcc gct gcg ttg caa gcc cag tac cac
gly his gln ser gly asp arg)AMB(ala arg   lys ala ala ala leu gln ala gln tyr his
541/181   SEQ ID NO: 801 ⌐   571/191
ctg ctg ttg cca cca ctg gcc ggg cgc ccc   ggg ata gcc gta cgc cac tcc gag cat tgg
leu leu leu pro pro leu ala gly arg pro   gly ile ala val arg his ser glu his trp
601/201   631/211
cgc gtt gct cag ttc ggc ggc cga cgg cag   cgc cgt ggt gtc ggc ggc ctc ggc ctg ttc
arg val ala gln phe gly gly arg arg gln   arg arg gly val gly gly leu gly leu phe
661/221   691/231
ggc tgc cgt tac ctc gac ggc cgc gac cgc   ctg cca gcc gcg ccg ccg gat gtg ctc cag
gly cys arg tyr leu asp gly arg asp arg   leu pro ala ala pro pro asp val leu gln
721/241   751/251
cca cat tgg ggc gcg caa agt ctc ggt gcc   cct ggg gta gcg cat cgc gtc gac ata cac
pro his trp gly ala gln ser leu gly ala   pro gly val ala his arg val asp ile his
781/261   811/271
cgt cag ggc atc acc gag gcg gcg ctc cat   atc gct ggg cgg cag atc gat gag gaa tat
arg gln gly ile thr glu ala ala leu his   ile ala gly arg gln ile asp glu glu tyr
841/281   871/291
cgc caa cgc gcg gtg tcc tcc tca tgt gat   gaa ccg atg cgt gct tgc gca cca gta tcg
arg gln arg ala val ser ser ser cys asp   glu pro met arg ala cys ala pro val ser
901/301   931/311
gac aag ccg atg agg ccg ccc gcg ctg gac   ggg gct tgt agc gta tgg ccg ttt ccg ctc
asp lys pro met arg pro pro ala leu asp   gly ala cys ser val trp pro phe pro leu

SEQ ID NOS:794-804

FIG. 45ZC

961/321
agc tcg tcg ctg cgg cgc cgc cgg gat aga
ser ser ser leu arg arg arg arg asp arg
1021/341
ttg acc tcg tat cat ctg agt tag ttg ccc
leu thr ser tyr his leu ser)AMB(leu pro
1081/361   SEQ ID NO: 802 ─◢
tac gtg aca gtc tgt cgg caa gga ggg acg
tyr val thr val cys arg gln gly gly thr
1141/381
gct tga cca gat cga gag cgc tct cta cgc
ala)OPA(pro asp arg glu arg ser leu arg
1201/401 ↖─ SEQ ID NO: 804
tgg cgg ggg ctt ccg cgc acc gac cgc gcg
trp arg gly leu pro arg thr asp arg ala
1261/421
cat cgg tct ggg gat gtt ggt ttc cgg cgt
his arg ser gly asp val gly phe arg arg
1321/441
ccc gat act cag cgt ttt cgg ttt tgt cgt
pro asp thr gln arg phe arg phe cys arg
1381/461
cgg tcc tcg gtt gtc cgg cag gat gga tcg
arg ser ser val val arg gln asp gly ser
1441/481
gcg tcg tac caa ggg ggc cgg ggg ctc att
ala ser tyr gln gly gly arg gly leu ile

991/331
atc gcc cgc gaa cca gtg gta cgg cgc aga
ile ala arg glu pro val val arg arg arg
1051/351
gcg caa tgg gca tcc gcg tgt tat cgg tat
ala gln trp ala ser ala cys tyr arg tyr
1111/371
cat gcc act ctc cga tca tga gca gcg gat
his ala thr leu arg ser)OPA(ala ala asp
1171/391   SEQ ID NO: 803 ─◢
cga aga tcc caa gtt cgc atc gag tgt ccg
arg arg ser gln val arg ile glu cys pro
1231/411
gcg gcg cct gca ggg cgc ggc gtt gtt cat
ala ala pro ala gly arg gly val val his
1291/431
ggc gtt caa aga gac cat gat cgg aag ttt
gly val gln arg asp his asp arg lys phe
1351/451
gat gtt cgg tgg tgt ggt gta tgc cat cac
asp val arg trp cys gly val cys his his
1411/471
tgg cgg atc ggc tgc tgg ggc ttc gcg cca
trp arg ile gly cys trp gly phe ala pro
1471/491
cac cag ccg tat gga aga tc
his gln pro tyr gly arg)

SEQ ID NOS:794-804 (continued 1)

FIG. 45ZC (continued 1)

SEQ ID NO: 805
⌐   seq 45ZA joined directly to phoA
cag tct gtc ggc aag gag gga cgc atg cca ctc tcc gat cat gag cag cgg
(gln ser val gly lys glu gly arg met pro leu ser asp his glu gln arg
1141/381 ─◥ SEQ ID NO: 806   1171/391
atg ctt gac cag atc gag agc gct ctc tac gcc gaa gat ccc aag ttc gca tcg agt gtc
met leu asp gln ile glu ser ala leu tyr ala glu asp pro lys phe ala ser ser val
1201/401   1231/411
cgt ggc ggg ggc ttc cgc gca ccg acc gcg cgg cgg cgc ctg cag ggc gcg gcg ttg ttc
arg gly gly gly phe arg ala pro thr ala arg arg arg leu gln gly ala ala leu phe
1261/421   1291/431
atc atc ggt ctg ggg atg ttg gtt tcc ggc gtg gcg ttc aaa gag acc atg atc gga agt
ile ile gly leu gly met leu val ser gly val ala phe lys glu thr met ile gly ser
1321/441   1351/451
ttc ccg ata ctc agc gtt ttc ggt ttt gtc gtg atg ttc ggt ggt gtg gtg tat gcc atc
phe pro ile leu ser val phe gly phe val val met phe gly gly val val tyr ala ile
1381/461   1411/471
acc ggt cct cgg ttg tcc ggc agg atg gat cgt ggc gga tcg gct gct ggg gct tcg cgc
thr gly pro arg leu ser gly arg met asp arg gly gly ser ala ala gly ala ser arg
1441/481   1471/491
cag cgt cgt acc aag ggg gcc ggg ggc tca ttc acc agc cgt atg gaa gat c
gln arg arg thr lys gly ala gly gly ser phe thr ser arg met glu asp)

SEQ ID NOS:805-806

FIG. 45A

Sequence Rv2169c predicted by Cole et al., 1998 (Nature 393:537-544)
and containing Seq45A

1/1 ⟋ **SEQ ID NO: 807**   31/11

```
atg cca ctc tcc gat cat gag cag cgg atg   ctt gac cag atc gag agc gct ctc tac gcc
(Met pro leu ser asp his glu gln arg met   leu asp gln ile glu ser ala leu tyr ala
```
61/21 ⟍ **SEQ ID NO: 808**   91/31

```
gaa gat ccc aag ttc gca tcg agt gtc cgt   ggc ggg ggc ttc cgc gca ccg acc gcg cgg
glu asp pro lys phe ala ser ser val arg   gly gly gly phe arg ala pro thr ala arg
```
121/41   151/51

```
cgg cgc ctg cag ggc gcg gcg ttg ttc atc   atc ggt ctg ggg atg ttg gtt tcc ggc gtg
arg arg leu gln gly ala ala leu phe ile   ile gly leu gly met leu val ser gly val
```
181/61   211/71

```
gcg ttc aaa gag acc atg atc gga agt ttc   ccg ata ctc agc gtt ttc ggt ttt gtc gtg
ala phe lys glu thr met ile gly ser phe   pro ile leu ser val phe gly phe val val
```
241/81   271/91

```
atg ttc ggt ggt gtg gtg tat gcc atc acc   ggt cct cgg ttg tcc ggc agg atg gat cgt
met phe gly gly val val tyr ala ile thr   gly pro arg leu ser gly arg met asp arg
```
301/101   331/111

```
ggc gga tcg gct gct ggg gct tcg cgc cag   cgt cgt acc aag ggg gcc ggg ggc tca ttc
gly gly ser ala ala gly ala ser arg gln   arg arg thr lys gly ala gly gly ser phe
```
361/121   391/131

```
acc agc cgt atg gaa gat cgg ttc cgg cgc   cgc ttc gac gag taa
thr ser arg met glu asp arg phe arg arg   arg phe asp glu)OCH
```

SEQ ID NOS: 807-808

FIG. 45D

ORF according to Cole et al., 1998 (Nature 393:537-544) and containing Rv2169c

1/1 ⟋ **SEQ ID NO: 809**   31/11

```
tga cag tct gtc ggc aag gag gga cgc atg   cca ctc tcc gat cat gag cag cgg atg ctt
OPA(gln ser val gly lys glu gly arg met   pro leu ser asp his glu gln arg met leu
```
61/21 ⟍ **SEQ ID NO: 810**   91/31

```
gac cag atc gag agc gct ctc tac gcc gaa   gat ccc aag ttc gca tcg agt gtc cgt ggc
asp gln ile glu ser ala leu tyr ala glu   asp pro lys phe ala ser ser val arg gly
```
121/41   151/51

```
ggg ggc ttc cgc gca ccg acc gcg cgg cgg   cgc ctg cag ggc gcg gcg ttg ttc atc atc
gly gly phe arg ala pro thr ala arg arg   arg leu gln gly ala ala leu phe ile ile
```
181/61   211/71

```
ggt ctg ggg atg ttg gtt tcc ggc gtg gcg   ttc aaa gag acc atg atc gga agt ttc ccg
gly leu gly met leu val ser gly val ala   phe lys glu thr met ile gly ser phe pro
```
241/81   271/91

```
ata ctc agc gtt ttc ggt ttt gtc gtg atg   ttc ggt ggt gtg gtg tat gcc atc acc ggt
ile leu ser val phe gly phe val val met   phe gly gly val val tyr ala ile thr gly
```
301/101   331/111

```
cct cgg ttg tcc ggc agg atg gat cgt ggc   gga tcg gct gct ggg gct tcg cgc cag cgt
pro arg leu ser gly arg met asp arg gly   gly ser ala ala gly ala ser arg gln arg
```
361/121   391/131

```
cgt acc aag ggg gcc ggg ggc tca ttc acc   agc cgt atg gaa gat cgg ttc cgg cgc cgc
arg thr lys gly ala gly gly ser phe thr   ser arg met glu asp arg phe arg arg arg
```
421/141

```
ttc gac gag taa
phe asp glu)OCH
```

SEQ ID No:809-810

FIG. 45E

1/1 SEQ ID NO: 811   31/11
cag ccg cgc cgc atc gac cag ggc ctc acg ccc ggt cac ttc tcc gcg ttc ctc aac aat
(gln pro arg arg ile asp gln gly leu thr pro gly his phe ser ala phe leu asn asn
61/21 SEQ ID NO: 812   91/31
tcc ggt gaa cat cgc acc agg tta ggc agc aat ccc gcg gac ccg cac ccc act cgc cga
ser gly glu his arg thr arg leu gly ser asn pro ala asp pro his pro thr arg arg
121/41   151/51
ccg gcc aac tca cag aca ccc tct acg atg cag ggt atg cgg acc ccc aga cgc cac tgc
pro ala asn ser gln thr pro ser thr met gln gly met arg thr pro arg arg his cys
181/61   211/71
cgt cgc atc gcc gtc ctc gcc gcc gtt agc atc gcc gcc act gtc gtt gcc ggc tgc tcg
arg arg ile ala val leu ala ala val ser ile ala ala thr val val ala gly cys ser
241/81   271/91
tcg ggc tcg aag cca agc ggc gga cca ctt ccg gac gcg aag ccg ctg gtc gag gag gcc
ser gly ser lys pro ser gly gly pro leu pro asp ala lys pro leu val glu glu ala
301/101   331/111
acc gcg cag acc aag gct ctc aag agc gcg cac atg gtg ctg acg gtc aac ggc aag atc
thr ala gln thr lys ala leu lys ser ala his met val leu thr val asn gly lys ile)

SEQ ID NOS:811-812

FIG. 46A

1/1 SEQ ID NO: 813   31/11
agc cgc gcc gca tcg acc agg gcc tca cgc ccg gtc act tct ccg cgt tcc tca aca att
(ser arg ala ala ser thr arg ala ser arg pro val thr ser pro arg ser ser thr ile
61/21 SEQ ID NO: 814   91/31
ccg gtg aac atc gca cca ggt tag gca gca atc ccg cgg acc cgc acc cca ctc gcc gac
pro val asn ile ala pro gly)AMB(ala ala ile pro arg thr arg thr pro leu ala asp
121/41   SEQ ID NO: 815   151/51
cgg cca act cac aga cac cct cta cga tgc agg gta tgc gga ccc cca gac gcc act gcc
arg pro thr his arg his pro leu arg cys arg val cys gly pro pro asp ala thr ala
181/61   211/71
gtc gca tcg ccg tcc tcg ccg ccg tta gca tcg ccg cca ctg tcg ttg ccg gct gct cgt
val ala ser pro ser ser pro pro leu ala ser pro pro leu ser leu pro ala ala arg
241/81   271/91
cgg gct cga agc caa gcg gcg gac cac ttc cgg acg cga agc cgc tgg tcg agg agg cca
arg ala arg ser gln ala ala asp his phe arg thr arg ser arg trp ser arg arg pro
301/101   331/111
ccg cgc aga cca agg ctc tca aga gcg cgc aca tgg tgc tga cgg tca acg gca aga tc
pro arg arg pro arg leu ser arg ala arg thr trp cys)OPA(arg ser thr ala arg)
                                            SEQ ID NO: 816

SEQ ID NOS:813-816

FIG. 46B

1/1 SEQ ID NO: 817   31/11
gcc gcg ccg cat cga cca ggg cct cac gcc cgg tca ctt ctc cgc gtt cct caa caa ttc
(ala ala pro his arg pro gly pro his ala arg ser leu leu arg val pro gln gln phe
61/21 SEQ ID NO: 818   91/31
cgg tga aca tcg cac cag gtt agg cag caa tcc cgc gga ccc gca ccc cac tcg ccg acc
arg)OPA(thr ser his gln val arg gln gln ser arg gly pro ala pro his ser pro thr
121/41 SEQ ID NO: 819   151/51
ggc caa ctc aca gac acc ctc tac gat gca ggg tat gcg gac ccc cag acg cca ctg ccg
gly gln leu thr asp thr leu tyr asp ala gly tyr ala asp pro gln thr pro leu pro
181/61   211/71
tcg cat cgc cgt cct cgc cgc cgt tag cat cgc cgc cac tgt cgt tgc cgg ctg ctc gtc
ser his arg arg pro arg arg arg)AMB(his arg arg his cys arg cys arg leu leu val
241/81   SEQ ID NO: 820   271/91
ggg ctc gaa gcc aag cgg cgg acc act tcc gga cgc gaa gcc gct ggt cga gga ggc cac
gly leu glu ala lys arg arg thr thr ser gly arg glu ala ala gly arg gly gly his
301/101   331/111
cgc gca gac caa ggc tct caa gag cgc gca cat ggt gct gac ggt caa cgg caa gat c
arg ala asp gln gly ser gln glu arg ala his gly ala asp gly gln arg gln asp)

SEQ ID NOS:817-820

FIG. 46C

Coding sequence Rv1411c predicted by Cole et al., 1998 (Nature 393: 537-544) and containing seq46A:

```
1/1    SEQ ID NO: 821                      31/11
atg cgg acc ccc aga cgc cac tgc cgt cgc atc gcc gtc ctc gcc gcc gtt agc atc gcc
(Met arg thr pro arg arg his cys arg arg ile ala val leu ala ala val ser ile ala
61/21    SEQ ID NO: 822                    91/31
gcc act gtc gtt gcc ggc tgc tcg tcg ggc tcg aag cca agc ggc gga cca ctt ccg gac
ala thr val val ala gly cys ser ser gly ser lys pro ser gly gly pro leu pro asp
121/41                                     151/51
gcg aag ccg ctg gtc gag gag gcc acc gcg cag acc aag gct ctc aag agc gcg cac atg
ala lys pro leu val glu glu ala thr ala gln thr lys ala leu lys ser ala his met
181/61                                     211/71
gtg ctg acg gtc aac ggc aag atc ccg gga ctg tct ctg aag acg ctg agc ggc gat ctc
val leu thr val asn gly lys ile pro gly leu ser leu lys thr leu ser gly asp leu
241/81                                     271/91
acc acc aac ccc acc gcc gcg acg gga aac gtc aag ctc acg ctg ggt ggg tct gat atc
thr thr asn pro thr ala ala thr gly asn val lys leu thr leu gly gly ser asp ile
301/101                                    331/111
gat gcc gac ttc gtg gtg ttc gac ggg atc ctg tac gcc acc ctg acg ccc aac cag tgg
asp ala asp phe val val phe asp gly ile leu tyr ala thr leu thr pro asn gln trp
361/121                                    391/131
agc gat ttc ggt ccc gcc gcc gac atc tac gac ccc gcc cag gtg ctg aat ccg gat acc
ser asp phe gly pro ala ala asp ile tyr asp pro ala gln val leu asn pro asp thr
421/141                                    451/151
ggc ctg gcc aac gtg ctg gcg aat ttc gcc gac gca aaa gcc gaa ggg cgg gat acc atc
gly leu ala asn val leu ala asn phe ala asp ala lys ala glu gly arg asp thr ile
481/161                                    511/171
aac ggc cag aac acc atc cgc atc agc ggg aag gta tcg gca cag gcg gtg aac cag ata
asn gly gln asn thr ile arg ile ser gly lys val ser ala gln ala val asn gln ile
541/181                                    571/191
gcg ccg ccg ttc aac gcg acg cag ccg gtg ccg gcg acc gtc tgg att cag gag acc ggc
ala pro pro phe asn ala thr gln pro val pro ala thr val trp ile gln glu thr gly
601/201                                    631/211
gat cat caa ctg gca cag gcc cag ttg gac cgc ggc tcg ggc aat tcc gtc cag atg acc
asp his gln leu ala gln ala gln leu asp arg gly ser gly asn ser val gln met thr
661/221                                    691/231
ttg tcg aaa tgg ggc gag aag gtc cag gtc acg aag ccc ccg gtg agc tga
leu ser lys trp gly glu lys val gln val thr lys pro pro val ser)OPA
```

SEQ ID NOS:821-822

FIG. 46D

ORF according to Cole et al., 1998 (Nature 393: 537-544):
and containing the coding sequence Rv1411c:

```
1/1   SEQ ID NO: 823                    31/11
tag ctc acc cag gtt gga ccg gtt cag tgt ctc ggc cat cac gtc ggc ggt gaa ttg gcc
AMB(leu thr gln val gly pro val gln cys leu gly his his val gly gly glu leu ala
61/21  SEQ ID NO: 824                   91/31
gtc ggg caa tac atc gac gac cgt cag aca cac gcc gtt gac agc gat cga gtc gcc gtg
val gly gln tyr ile asp asp arg gln thr his ala val asp ser asp arg val ala val
121/41                                  151/51
gcc ggc gtc ggc ggt aac cat cgg acc gcg gat ggt cag ccg cgc cgc atc gac cag ggc
ala gly val gly gly asn his arg thr ala asp gly gln pro arg arg ile asp gln gly
181/61                                  211/71
ctc acg ccc ggt cac ttc tcc gcg ttc ctc aac aat tcc ggt gaa cat cgc acc agg tta
leu thr pro gly his phe ser ala phe leu asn asn ser gly glu his arg thr arg leu
241/81                                  271/91
ggc agc aat ccc gcg gac ccg cac ccc act cgc cga ccg gcc aac tca cag aca ccc tct
gly ser asn pro ala asp pro his pro thr arg arg pro ala asn ser gln thr pro ser
301/101                                 331/111
acg atg cag ggt atg cgg acc ccc aga cgc cac tgc cgt cgc atc gcc gtc ctc gcc gcc
thr met gln gly met arg thr pro arg arg his cys arg arg ile ala val leu ala ala
361/121                                 391/131
gtt agc atc gcc gcc act gtc gtt gcc ggc tgc tcg tcg ggc tcg aag cca agc ggc gga
val ser ile ala ala thr val val ala gly cys ser ser gly ser lys pro ser gly gly
421/141                                 451/151
cca ctt ccg gac gcg aag ccg ctg gtc gag gag gcc acc gcg cag acc aag gct ctc aag
pro leu pro asp ala lys pro leu val glu glu ala thr ala gln thr lys ala leu lys
481/161                                 511/171
agc gcg cac atg gtg ctg acg gtc aac ggc aag atc ccg gga ctg tct ctg aag acg ctg
ser ala his met val leu thr val asn gly lys ile pro gly leu ser leu lys thr leu
541/181                                 571/191
agc ggc gat ctc acc acc aac ccc acc gcc gcg acg gga aac gtc aag ctc acg ctg ggt
ser gly asp leu thr thr asn pro thr ala ala thr gly asn val lys leu thr leu gly
601/201                                 631/211
ggg tct gat atc gat gcc gac ttc gtg gtg ttc gac ggg atc ctg tac gcc acc ctg acg
gly ser asp ile asp ala asp phe val val phe asp gly ile leu tyr ala thr leu thr
661/221                                 691/231
ccc aac cag tgg agc gat ttc ggt ccc gcc gcc gac atc tac gac ccc gcc cag gtg ctg
pro asn gln trp ser asp phe gly pro ala ala asp ile tyr asp pro ala gln val leu
721/241                                 751/251
aat ccg gat acc ggc ctg gcc aac gtg ctg gcg aat ttc gcc gac gca aaa gcc gaa ggg
asn pro asp thr gly leu ala asn val leu ala asn phe ala asp ala lys ala glu gly
781/261                                 811/271
cgg gat acc atc aac ggc cag aac acc atc cgc atc agc ggg aag gta tcg gca cag gcg
arg asp thr ile asn gly gln asn thr ile arg ile ser gly lys val ser ala gln ala
841/281                                 871/291
gtg aac cag ata gcg ccg ccg ttc aac gcg acg cag ccg gtg ccg gcg acc gtc tgg att
val asn gln ile ala pro pro phe asn ala thr gln pro val pro ala thr val trp ile
901/301                                 931/311
cag gag acc ggc gat cat caa ctg gca cag gcc cag ttg gac cgc ggc tcg ggc aat tcc
gln glu thr gly asp his gln leu ala gln ala gln leu asp arg gly ser gly asn ser
961/321                                 991/331
gtc cag atg acc ttg tcg aaa tgg ggc gag aag gtc cag gtc acg aag ccc ccg gtg agc
val gln met thr leu ser lys trp gly glu lys val gln val thr lys pro pro val ser)
1021/341
tga
OPA
```

SEQ ID NOS:823-824

FIG. 46E

1/1 ⌐SEQ ID NO: 825   31/11
gag ctg gtc aac ggc gcc ggc atc gac gac gcc gcc gtc gtg acc tgc cgg ccg gac agc
(glu leu val asn gly ala gly ile asp asp ala ala val val thr cys arg pro asp ser
61/21 ⌐SEQ ID NO: 826   91/31
ctg gcc gat gcc cag cag atg gtc gag gcg gca ctg ggc cga tat ggc cgt ttg gac gga
leu ala asp ala gln gln met val glu ala ala leu gly arg tyr gly arg leu asp gly
121/41   151/51
gtg ttg gtg gcc tcg ggc agc aac cat gtg gcg ccc att acc gag atg gcc gtc gag gac
val leu val ala ser gly ser asn his val ala pro ile thr glu met ala val glu asp
181/61   211/71
ttc gac gct gtg atg gac gcg aac gtg cgg ggt gcc tgg ctg gtg tgt cgg gcg gcc gga
phe asp ala val met asp ala asn val arg gly ala trp leu val cys arg ala ala gly
241/81   271/91
cgg gtg ctg ctc gag cag ggt cag ggc ggc agc gtg gtg ctg gtg tcg tcc gtt cgc ggc
arg val leu leu glu gln gly gln gly gly ser val val leu val ser ser val arg gly
301/101   331/111
ggg ttg ggc aat gcc gcc ggt tac agc gcg tac tgc ccg tcg aag gcg ggc acc gat c
gly leu gly asn ala ala gly tyr ser ala tyr cys pro ser lys ala gly thr asp)

SEQ ID NOS: 825-826

FIG. 47A

1/1 ⌐SEQ ID NO: 827   31/11
agc tgg tca acg gcg ccg gca tcg acg acg ccg ccg tcg tga cct gcc ggc cgg aca gcc
(ser trp ser thr ala pro ala ser thr thr pro pro ser)OPA(pro ala gly arg thr ala
61/21 ⌐SEQ ID NO: 828   91/31   ⌐SEQ ID NO: 829
tgg ccg atg ccc agc aga tgg tcg agg cgg cac tgg gcc gat atg gcc gtt tgg acg gag
trp pro met pro ser arg trp ser arg arg his trp ala asp met ala val trp thr glu
121/41   151/51
tgt tgg tgg cct cgg gca gca acc atg tgg cgc cca tta ccg aga tgg ccg tcg agg act
cys trp trp pro arg ala ala thr met trp arg pro leu pro arg trp pro ser arg thr
181/61   211/71
tcg acg ctg tga tgg acg cga acg tgc ggg gtg cct ggc tgg tgt gtc ggg cgg ccg gac
ser thr leu)OPA(trp thr arg thr cys gly val pro gly trp cys val gly arg pro asp
241/81   ⌐SEQ ID NO: 830   271/91
ggg tgc tgc tcg agc agg gtc agg gcg gca gcg tgg tgc tgg tgt cgt ccg ttc gcg gcg
gly cys cys ser ser arg val arg ala ala ala trp cys trp cys arg pro phe ala ala
301/101   331/111
ggt tgg gca atg ccg ccg gtt aca gcg cgt act gcc gtc gaa ggc ggg cac cga tc
gly trp ala met pro pro val thr ala arg thr ala arg arg arg arg ala pro ile)

SEQ ID NOS: 827-830

FIG. 47B

1/1 SEQ ID NO: 831   31/11
gct ggt caa cgg cgc cgg cat cga cga cgc cgc cgt cgt gac ctg ccg gcc gga cag cct
(ala gly gln arg arg arg his arg arg arg arg arg arg asp leu pro ala gly gln pro
61/21 ──SEQ ID NO: 832   91/31
ggc cga tgc cca gca gat ggt cga ggc ggc act ggg ccg ata tgg ccg ttt gga cgg agt
gly arg cys pro ala asp gly arg gly gly thr gly pro ile trp pro phe gly arg ser
121/41   151/51
gtt ggt ggc ctc ggg cag caa cca tgt ggc gcc cat tac cga gat ggc cgt cga gga ctt
val gly gly leu gly gln gln pro cys gly ala his tyr arg asp gly arg arg gly leu
181/61   211/71
cga cgc tgt gat gga cgc gaa cgt gcg ggg tgc ctg gct ggt gtg tcg ggc ggc cgg acg
arg arg cys asp gly arg glu arg ala gly cys leu ala gly val ser gly gly arg thr
241/81   271/91
ggt gct gct cga gca ggg tca ggg cgg cag cgt ggt gct ggt gtc gtc cgt tcg cgg cgg
gly ala ala arg ala gly ser gly arg gln arg gly ala gly val val arg ser arg arg
301/101   331/111
gtt ggg caa tgc cgc cgg tta cag cgc gta ctg ccc gtc gaa ggc ggg cac cga tc
val gly gln cys arg arg leu gln arg val leu pro val glu gly gly his arg)

SEQ ID NOS:831-832

FIG. 47C

Coding sequence Rv1714 predicted by Cole et al., 1998 (Nature 393: 537-544) and containing seq 47A:

```
1/1   SEQ ID NO: 833                    31/11
gtg gag gaa atg gcg ctg gct cag cag gtg ccg aac ctg ggt ctg gcg cgc ttc agc gtg
(val glu glu met ala leu ala gln gln val pro asn leu gly leu ala arg phe ser val
61/21   SEQ ID NO: 834                  91/31
cag gac aag tcg atc ctg atc acc ggc gcg acc ggt tcg ttg ggc cga gtt gcc gcc cgg
gln asp lys ser ile leu ile thr gly ala thr gly ser leu gly arg val ala ala arg
121/41                                  151/51
gcg ctg gcc gac gcg gga gcg cgg ctg aca ctg gcc ggc ggc aac tcg gcc ggt ctg gcc
ala leu ala asp ala gly ala arg leu thr leu ala gly gly asn ser ala gly leu ala
181/61                                  211/71
gag ctg gtc aac ggc gcc ggc atc gac gac gcc gcc gtc gtg acc tgc cgg ccg gac agc
glu leu val asn gly ala gly ile asp asp ala ala val val thr cys arg pro asp ser
241/81                                  271/91
ctg gcc gat gcc cag cag atg gtc gag gcg gca ctg ggc cga tat ggc cgt ttg gac gga
leu ala asp ala gln gln met val glu ala ala leu gly arg tyr gly arg leu asp gly
301/101                                 331/111
gtg ttg gtg gcc tcg ggc agc aac cat gtg gcg ccc att acc gag atg gcc gtc gag gac
val leu val ala ser gly ser asn his val ala pro ile thr glu met ala val glu asp
361/121                                 391/131
ttc gac gct gtg atg gac gcg aac gtg cgg ggt gcc tgg ctg gtg tgt cgg gcg gcc gga
phe asp ala val met asp ala asn val arg gly ala trp leu val cys arg ala ala gly
421/141                                 451/151
cgg gtg ctg ctc gag cag ggt cag ggc ggc agc gtg gtg ctg gtg tcg tcc gtt cgc ggc
arg val leu leu glu gln gly gln gly gly ser val val leu val ser ser val arg gly
481/161                                 511/171
ggg ttg ggc aat gcc gcc ggt tac agc gcg tac tgc ccg tcg aag gcg ggc acc gat ctg
gly leu gly asn ala ala gly tyr ser ala tyr cys pro ser lys ala gly thr asp leu
541/181                                 571/191
ttg gcc aag aca ttg gcg gcc gaa tgg ggc ggt cac ggc att cgg gtg aac gcg ctg gcg
leu ala lys thr leu ala ala glu trp gly gly his gly ile arg val asn ala leu ala
601/201                                 631/211
ccg acg gtg ttt cgg tcc gcg gtg acc gag tgg atg ttc acc gac gat ccg aag ggc cgg
pro thr val phe arg ser ala val thr glu trp met phe thr asp asp pro lys gly arg
661/221                                 691/231
gcc acc cgg gag gcg atg ctc gcc cgg atc ccg ttg cgc cgc ttc gcc gaa ccg gaa gac
ala thr arg glu ala met leu ala arg ile pro leu arg arg phe ala glu pro glu asp
721/241                                 751/251
ttc gtc ggc gcc ctg atc tat ctg ctc agc gac gcc tcg agc ttc tac acc ggc cag gtg
phe val gly ala leu ile tyr leu leu ser asp ala ser ser phe tyr thr gly gln val
781/261                                 811/271
atg tat ctg gac ggc ggg tac acc gca tgc tga
met tyr leu asp gly gly tyr thr ala cys)OPA
```

SEQ ID NOS:833-834

FIG. 47D

ORF according to Cole et al., 1998 (Nature 393: 537-544) and containing
the coding sequence Rv1714:

24/1 SEQ ID NO: 835   54/11
tag gtg gag gaa atg gcg ctg gct cag cag gtg ccg aac ctg ggt ctg gcg cgc ttc agc
AMB(val glu glu met ala leu ala gln gln val pro asn leu gly leu ala arg phe ser
84/21 SEQ ID NO: 836   114/31
gtg cag gac aag tcg atc ctg atc acc ggc gcg acc ggt tcg ttg ggc cga gtt gcc gcc
val gln asp lys ser ile leu ile thr gly ala thr gly ser leu gly arg val ala ala
144/41   174/51
cgg gcg ctg gcc gac gcg gga gcg cgg ctg aca ctg gcc ggc ggc aac tcg gcc ggt ctg
arg ala leu ala asp ala gly ala arg leu thr leu ala gly gly asn ser ala gly leu
204/61   234/71
gcc gag ctg gtc aac ggc gcc ggc atc gac gac gcc gcc gtc gtg acc tgc cgg ccg gac
ala glu leu val asn gly ala gly ile asp asp ala ala val val thr cys arg pro asp
264/81   294/91
agc ctg gcc gat gcc cag cag atg gtc gag gcg gca ctg ggc cga tat ggc cgt ttg gac
ser leu ala asp ala gln gln met val glu ala ala leu gly arg tyr gly arg leu asp
324/101   354/111
gga gtg ttg gtg gcc tcg ggc agc aac cat gtg gcg ccc att acc gag atg gcc gtc gag
gly val leu val ala ser gly ser asn his val ala pro ile thr glu met ala val glu
384/121   414/131
gac ttc gac gct gtg atg gac gcg aac gtg cgg ggt gcc tgg ctg gtg tgt cgg gcg gcc
asp phe asp ala val met asp ala asn val arg gly ala trp leu val cys arg ala ala
444/141   474/151
gga cgg gtg ctg ctc gag cag ggt cag ggc ggc agc gtg gtg ctg gtg tcg tcc gtt cgc
gly arg val leu leu glu gln gly gln gly gly ser val val leu val ser ser val arg
504/161   534/171
ggc ggg ttg ggc aat gcc gcc ggt tac agc gcg tac tgc ccg tcg aag gcg ggc acc gat
gly gly leu gly asn ala ala gly tyr ser ala tyr cys pro ser lys ala gly thr asp
564/181   594/191
ctg ttg gcc aag aca ttg gcg gcc gaa tgg ggc ggt cac ggc att cgg gtg aac gcg ctg
leu leu ala lys thr leu ala ala glu trp gly gly his gly ile arg val asn ala leu
624/201   654/211
gcg ccg acg gtg ttt cgg tcc gcg gtg acc gag tgg atg ttc acc gac gat ccg aag ggc
ala pro thr val phe arg ser ala val thr glu trp met phe thr asp asp pro lys gly
684/221   714/231
cgg gcc acc cgg gag gcg atg ctc gcc cgg atc ccg ttg cgc cgc ttc gcc gaa ccg gaa
arg ala thr arg glu ala met leu ala arg ile pro leu arg arg phe ala glu pro glu
744/241   774/251
gac ttc gtc ggc gcc ctg atc tat ctg ctc agc gac gcc tcg agc ttc tac acc ggc cag
asp phe val gly ala leu ile tyr leu leu ser asp ala ser ser phe tyr thr gly gln
804/261   834/271
gtg atg tat ctg gac ggc ggg tac acc gca tgc tga
val met tyr leu asp gly gly tyr thr ala cys)OPA

SEQ ID NOS:835-836

FIG. 47E

```
1/1    SEQ ID NO: 837                          31/11
agg ctc atg agc aag acg gtt ctc atc ctt ggc gcg ggt gtc ggc ggc ctg acc acc gcc
(arg leu met ser lys thr val leu ile leu gly ala gly val gly gly leu thr thr ala
61/21   SEQ ID NO: 838                          91/31
gac acc ctc cgt caa ctg cta cca cct gag gat c
asp thr leu arg gln leu leu pro pro glu asp)
```

SEQ ID NOS:837-838

FIG. 48A

```
1/1    SEQ ID NO: 839                          31/11
ggc tca tga gca aga cgg ttc tca tcc ttg gcg cgg gtg tcg gcg gcc tga cca ccg ccg
gly ser OPA(ala arg arg phe ser ser leu ala arg val ser ala ala)OPA(pro pro pro
61/21           SEQ ID NO: 840                 91/31        SEQ ID NO: 841
aca ccc tcc gtc aac tgc tac cac ctg agg atc
thr pro ser val asn cys tyr his leu arg ile)
```

SEQ ID NOS:839-841

FIG. 48B

```
1/1    SEQ ID NO: 842                          31/11
gct cat gag caa gac ggt tct cat cct tgg cgc ggg tgt cgg cgg cct gac cac cgc cga
(ala his glu gln asp gly ser his pro trp arg gly cys arg arg pro asp his arg arg
61/21   SEQ ID NO: 843
cac cct ccg tca act gct acc acc tga gga tc
his pro pro ser thr ala thr thr)OPA gly
```

SEQ ID NOS:842-843

FIG. 48C

Coding sequence Rv0331 predicted by Cole et al., 1998 (Nature 393: 537-544) and containing seq48A:

```
1/1  SEQ ID NO: 844                 31/11
atg agc aag acg gtt ctc atc ctt ggc gcg ggt gtc ggc ggc ctg acc acc gcc gac acc
(Met ser lys thr val leu ile leu gly ala gly val gly gly leu thr thr ala asp thr
 61/21  -SEQ ID NO: 845            91/31
ctc cgt caa ctg cta cca cct gag gat cga atc ata ttg gtg gac agg agc ttt gac ggg
leu arg gln leu leu pro pro glu asp arg ile ile leu val asp arg ser phe asp gly
121/41                            151/51
acg ctg ggc ttg tcg ttg cta tgg gtg ttg cgg ggc tgg cgg cgg cct gac gac gtc cgc
thr leu gly leu ser leu leu trp val leu arg gly trp arg arg pro asp asp val arg
181/61                            211/71
gtc cgc ccc acc gcg gcg tcg ctg ccc ggt gtg gaa atg gtt act gca acc gtc gcc cac
val arg pro thr ala ala ser leu pro gly val glu met val thr ala thr val ala his
241/81                            271/91
att gac atc gcg gcc cag gta gtg cac acc gac aac agc gtc atc ggc tat gac gcg ttg
ile asp ile ala ala gln val val his thr asp asn ser val ile gly tyr asp ala leu
301/101                           331/111
gtg atc gca tta ggt gcg gcg ctg aac acc gac gcc gtt ccc gga ctg tcg gac gcg ctc
val ile ala leu gly ala ala leu asn thr asp ala val pro gly leu ser asp ala leu
361/121                           391/131
gac gcc gac gtc gcg ggc cag ttc tac acc ctg gac ggc gcg gct gag ctg cgt gcg aag
asp ala asp val ala gly gln phe tyr thr leu asp gly ala ala glu leu arg ala lys
421/141                           451/151
gtc gag gcg ctc gag cat ggc cgg atc gct gtg gct atc gcc ggg gtg ccg ttc aaa tgc
val glu ala leu glu his gly arg ile ala val ala ile ala gly val pro phe lys cys
481/161                           511/171
cca gcc gca ccg ttc gaa gcg gcg ttt ctg atc gcc gcc caa ctc ggt gac cgc tac gcc
pro ala ala pro phe glu ala ala phe leu ile ala ala gln leu gly asp arg tyr ala
541/181                           571/191
acc gga acc gta cag atc gac acg ttc acg cct gac ccg ctg ccg atg ccc gtt gca ggt
thr gly thr val gln ile asp thr phe thr pro asp pro leu pro met pro val ala gly
601/201                           631/211
ccc gag gtc ggc gag gct ttg gtc tcg atg ctc aag gat cac ggt gtc ggc ttc cat cct
pro glu val gly glu ala leu val ser met leu lys asp his gly val gly phe his pro
661/221                           691/231
cgc aag gcc cta gct cgc gtc gat gag gcc gca agg acg atg cac ttc ggt gac ggc acg
arg lys ala leu ala arg val asp glu ala ala arg thr met his phe gly asp gly thr
721/241                           751/251
tcc gaa ccg ttc gat ctg ctt gcc gtg gtc ccc ccg cac gtg ccc tcc gcc gcg gcg cgg
ser glu pro phe asp leu leu ala val val pro pro his val pro ser ala ala ala arg
781/261                           811/271
tca gcg ggt ctc agc gaa tcc ggg tgg ata ccc gtg gac ccg cgc acc ctg tcc act agc
ser ala gly leu ser glu ser gly trp ile pro val asp pro arg thr leu ser thr ser
841/281                           871/291
gcc gac aac gtg tgg gcc atc ggc gat gcg acc gtg ctg acg ctg ccg aat ggc aaa ccg
ala asp asn val trp ala ile gly asp ala thr val leu thr leu pro asn gly lys pro
901/301                           931/311
ctg ccc aag gct gcc gtg ttc gcc gaa gcc cag gcc gca gtt gtc gcc cac ggc gtc gcc
leu pro lys ala ala val phe ala glu ala gln ala ala val val ala his gly val ala
961/321                           991/331
cgc cat ctc ggt tac gac gta gct gag cgc cac ttc acc ggc acg ggc gcc tgc tac gtc
arg his leu gly tyr asp val ala glu arg his phe thr gly thr gly ala cys tyr val
1021/341                          1051/351
gag acc ggt gat cac cag gca gcc aag ggc gac ggc gat ttc ttc gct ccg tcg gcg ccc
glu thr gly asp his gln ala ala lys gly asp gly asp phe phe ala pro ser ala pro
1081/361                          1111/371
tcg gtg acg ctg tac ccg ccg tcg cgg gag ttt cac gag gag aag gtc gca caa gaa ctg
ser val thr leu tyr pro pro ser arg glu phe his glu glu lys val ala gln glu leu
1141/381
gcc tgg ctg acc cgc tgg aag acg tga
ala trp leu thr arg trp lys thr)OPA
```

SEQ ID NOS:844-845

FIG. 48D

ORF according to Cole et al., 1998 (Nature 393: 537-544) and containing
coding sequence Rv0331:

```
1/1  SEQ ID NO: 846                    31/11
tga aca ccc gcg ccg acg cgg cga caa tcg cgg aaa acc ggt ccg cgg gaa tgc tgc ggg
OPA(thr pro ala pro thr arg arg gln ser arg lys thr gly pro arg glu cys cys gly
61/21   SEQ ID NO: 847                 91/31
cca tgg gcc gat aat agt ttg act gac tcg gtc agt cac ccc aag acc ttg cgc aag act
pro trp ala asp asn ser leu thr asp ser val ser his pro lys thr leu arg lys thr
121/41                                 151/51
gcg gcg gaa tct aat att cca aag ata tat gga act cga tgc gaa gga atc agg ctc atg
ala ala glu ser asn ile pro lys ile tyr gly thr arg cys glu gly ile arg leu met
181/61                                 211/71
agc aag acg gtt ctc atc ctt ggc gcg ggt gtc ggc ggc ctg acc acc gcc gac acc ctc
ser lys thr val leu ile leu gly ala gly val gly gly leu thr thr ala asp thr leu
241/81                                 271/91
cgt caa ctg cta cca cct gag gat cga atc ata ttg gtg gac agg agc ttt gac ggg acg
arg gln leu leu pro pro glu asp arg ile ile leu val asp arg ser phe asp gly thr
301/101                                331/111
ctg ggc ttg tcg ttg cta tgg gtg ttg cgg ggc tgg cgg cgg cct gac gac gtc cgc gtc
leu gly leu ser leu leu trp val leu arg gly trp arg arg pro asp asp val arg val
361/121                                391/131
cgc ccc acc gcg gcg tcg ctg ccc ggt gtg gaa atg gtt act gca acc gtc gcc cac att
arg pro thr ala ala ser leu pro gly val glu met val thr ala thr val ala his ile
421/141                                451/151
gac atc gcg gcc cag gta gtg cac acc gac aac agc gtc atc ggc tat gac gcg ttg gtg
asp ile ala ala gln val val his thr asp asn ser val ile gly tyr asp ala leu val
481/161                                511/171
atc gca tta ggt gcg gcg ctg aac acc gac gcc gtt ccc gga ctg tcg gac gcg ctc gac
ile ala leu gly ala ala leu asn thr asp ala val pro gly leu ser asp ala leu asp
541/181                                571/191
gcc gac gtc gcg ggc cag ttc tac acc ctg gac ggc gcg gct gag ctg cgt gcg aag gtc
ala asp val ala gly gln phe tyr thr leu asp gly ala ala glu leu arg ala lys val
601/201                                631/211
gag gcg ctc gag cat ggc cgg atc gct gtg gct atc gcc ggg gtg ccg ttc aaa tgc cca
glu ala leu glu his gly arg ile ala val ala ile ala gly val pro phe lys cys pro
661/221                                691/231
gcc gca ccg ttc gaa gcg gcg ttt ctg atc gcc gcc caa ctc ggt gac cgc tac gcc acc
ala ala pro phe glu ala ala phe leu ile ala ala gln leu gly asp arg tyr ala thr
721/241                                751/251
gga acc gta cag atc gac acg ttc acg cct gac ccg ctg ccg atg ccc gtt gca ggt ccc
gly thr val gln ile asp thr phe thr pro asp pro leu pro met pro val ala gly pro
781/261                                811/271
gag gtc ggc gag gct ttg gtc tcg atg ctc aag gat cac ggt gtc ggc ttc cat cct cgc
glu val gly glu ala leu val ser met leu lys asp his gly val gly phe his pro arg
841/281                                871/291
aag gcc cta gct cgc gtc gat gag gcc gca agg acg atg cac ttc ggt gac ggc acg tcc
lys ala leu ala arg val asp glu ala ala arg thr met his phe gly asp gly thr ser
```

SEQ ID NOS:846-847

FIG. 48E

901/301
gaa ccg ttc gat ctg ctt gcc gtg gtc ccc
glu pro phe asp leu leu ala val val pro
961/321
gcg ggt ctc agc gaa tcc ggg tgg ata ccc
ala gly leu ser glu ser gly trp ile pro
1021/341
gac aac gtg tgg gcc atc ggc gat gcg acc
asp asn val trp ala ile gly asp ala thr
1081/361
ccc aag gct gcc gtg ttc gcc gaa gcc cag
pro lys ala ala val phe ala glu ala gln
1141/381
cat ctc ggt tac gac gta gct gag cgc cac
his leu gly tyr asp val ala glu arg his
1201/401
acc ggt gat cac cag gca gcc aag ggc gac
thr gly asp his gln ala ala lys gly asp
1261/421
gtg acg ctg tac ccg ccg tcg cgg gag ttt
val thr leu tyr pro pro ser arg glu phe
1321/441
tgg ctg acc cgc tgg aag acg tga
trp leu thr arg trp lys thr)OPA

931/311
ccg cac gtg ccc tcc gcc gcg gcg cgg tca
pro his val pro ser ala ala ala arg ser
991/331
gtg gac ccg cgc acc ctg tcc act agc gcc
val asp pro arg thr leu ser thr ser ala
1051/351
gtg ctg acg ctg ccg aat ggc aaa ccg ctg
val leu thr leu pro asn gly lys pro leu
1111/371
gcc gca gtt gtc gcc cac ggc gtc gcc cgc
ala ala val val ala his gly val ala arg
1171/391
ttc acc ggc acg ggc gcc tgc tac gtc gag
phe thr gly thr gly ala cys tyr val glu
1231/411
ggc gat ttc ttc gct ccg tcg gcg ccc tcg
gly asp phe phe ala pro ser ala pro ser
1291/431
cac gag gag aag gtc gca caa gaa ctg gcc
his glu glu lys val ala gln glu leu ala

SEQ ID NOS:846-847 (continued)

FIG. 48E (continued)

Fragment amplified by PCR based on the sequence similarities with a
serine protease of the E.coli htrA family (creation of the BamHI site at
the 5' end and of the SnaBI site at the 3' end) and subcloned into the
vector pJVED:

```
1/1    SEQ ID NO: 848                              31/11
cca tct aca ccg ctc aac agc cgg gcc aga cgc tgc cgg tcg gtg ctg ccg aga agg cgg
(pro ser thr pro leu asn ser arg ala arg arg cys arg ser val leu pro arg arg arg)
61/21   SEQ ID NO: 849                             91/31
tga tcc gtg gcg agt tgt tca tgt cgc ggc gca cca ccg ccg acc aac ggg tgc ttg cca
OPA(ser val ala ser cys ser cys arg gly ala pro pro pro thr asn gly cys leu pro
121/41   SEQ ID NO: 850   SEQ ID NO: 852          151/51
tcc gtc tga cca acg gta gtt cgc tgc tga tct cca aaa gtc tca agc cca ccg aag cag
ser val)OPA(pro thr val val arg cys)OPA(ser pro lys val ser ser pro pro lys gln
181/61        SEQ ID NO: 851                       211/71
tca tga aca agc tgc gtt ggg tgc tat tga tcg tgg gtg gga tcg ggg tgg cgg tcg ccg
ser)OPA(thr ser cys val gly cys tyr)OPA(ser trp val gly ser gly trp arg ser pro
241/81   SEQ ID NO: 853                            271/91  SEQ ID NO: 854
cgg tgg ccg ggg gga tgg tca ccc ggg ccg ggc tga ggc cgg tgg gcc gcc tca ccg aag
arg trp pro gly gly trp ser pro gly pro gly)OPA(gly arg trp ala ala ser pro lys
301/101                                            331/111  SEQ ID NO: 855
cgg ccg agc ggg tgg cgc gaa ccg acg acc tgc ggc cca tcc ccg tct tcg gca gcg acg
arg pro ser gly trp arg glu pro thr thr cys gly pro ser pro ser ser ala ala thr
361/121                                            391/131
aat tgg cca ggc tga cag agg cat tca att taa tgc tgc ggg cgc tgg ccg agt cac ggg
asn trp pro gly)OPA(gln arg his ser ile)OCH(cys cys gly arg trp pro ser his gly
421/141        SEQ ID NO: 856                      451/151  SEQ ID NO: 857
aac ggc agg caa ggc tgg tta ccg acg ccg gac atg aat tgc gta ccc cgc taa cgt cgc
asn gly arg gln gly trp leu pro thr pro asp met asn cys val pro arg)OCH(arg arg
481/161                                            511/171       SEQ ID NO: 858
tgc gca cca atg tcg aac tct tga tgg cct cga tgg ccc cgg ggg ctc cgc ggc tac cca
cys ala pro met ser asn ser)OPA(trp pro arg trp pro arg gly leu arg gly tyr pro
541/181        SEQ ID NO: 859                      571/191
agc agg aga tgg tcg acc tgc gtg ccg atg tgc tgg ctc aaa tcg agg aat tgt cca cac
ser arg arg trp ser thr cys val pro met cys trp leu lys ser arg asn cys pro his
601/201                                            631/211
tgg tag gcg att tgg tgg acc tgt ccc gag gcg acg ccg gag aag tgg tgc acg agc cgg
trp)AMB(ala ile trp trp thr cys pro glu ala thr pro glu lys trp cys thr ser arg
661/221   SEQ ID NO: 860                           691/231
tcg aca tgg ctg acg tcg tcg acc gca gcc tgg agc ggg tca ggc ggc ggc gca acg ata
ser thr trp leu thr ser ser thr ala ala trp ser gly ser gly gly gly ala thr ile
721/241                                            751/251
tcc ttt tcg acg tcg agg tga ttg ggt ggc agg ttt atg gcg ata ccg ctg gat tgt cgc
ser phe ser thr ser arg)OPA(leu gly gly arg phe met ala ile pro leu asp cys arg
781/261   SEQ ID NO: 861                           811/271
gga tgg cgc tta acc tga tgg aca acg ccg cga agt gga gcc cgc cgg gcg gcc acg tgg
gly trp arg leu thr)OPA(trp thr thr pro arg ser gly ala arg arg ala ala thr trp
841/281   SEQ ID NO: 862                           871/291
gtg tca ggc tga gcc agc tcg acg cgt cgc acg ctg agc tgg tgg ttt ccg acc gcg gcc
val ser gly)OPA(ala ser ser thr arg arg thr leu ser trp trp phe pro thr ala ala
          SEQ ID NO: 863
```

                    SEQ ID NOS:848-863

FIG. 49A

901/301
cgg gca ttc ccg tgc agg agc gcc gtc tgg
arg ala phe pro cys arg ser ala val trp
961/321
cac ggg cgt tgc cgg gtt cgg gcc tcg ggt
his gly arg cys arg val arg ala ser gly
1021/341
acg gcg gat tgc tgc gca tcg aag aca ccg
thr ala asp cys cys ala ser lys thr pro
1081/361
ttt acg tgc tgc tcc ccg gcc gtc gga tgc
phe thr cys cys ser pro ala val gly cys
1141/381
gcg ctc gga gca cgg aca tcg aga act ctc
ala leu gly ala arg thr ser arg thr leu
1201/401
ctc agt cca cgc gcg caa cct agt tgt gca
leu ser pro arg ala gln pro ser cys ala
1261/421
cca cgc atg gcc aag ttg gcc cga gta gtg
pro arg met ala lys leu ala arg val val
1321/441
atg acg aat cac cca cgg tat tcg cca ccg
met thr asn his pro arg tyr ser pro pro
1381/461
cag ggg cag cag caa acg tac agc cag cag
gln gly gln gln gln thr tyr ser gln gln
1441/481
ccg cag cca acc cag tac cgt caa ccc tac
pro gln pro thr gln tyr arg gln pro tyr
1501/501
ata cct ggc gtg att ccg acc atg acg ccc
ile pro gly val ile pro thr met thr pro
1561/521
gca ggc atg ttg gcc atc ggc gcg gtg acg
ala gly met leu ala ile gly ala val thr
1621/541
gcg gcc gca tcc ctg gtc ggg ttc aac cgg
ala ala ala ser leu val gly phe asn arg
1681/561
gct gcc agc gcg gcg cca agc atc ccc gca
ala ala ser ala ala pro ser ile pro ala
1741/581
gtg gcg gcc aag gtg gtg ccc agt gtc gtc
val ala ala lys val val pro ser val val
1801/601
gag gag ggc tcc ggc atc att ctg tct gcc
glu glu gly ser gly ile ile leu ser ala
1861/621
atc gcg gcg gcc gcc aag cct ccc ctg ggc
ile ala ala ala ala lys pro pro leu gly

931/311
tgt ttg aac ggt ttt acc ggt cgg cat cgg
cys leu asn gly phe thr gly arg his arg
991/331
tgg cga tcg tca aac agg tgg tgc tca acc
trp arg ser ser asn arg trp cys ser thr
1051/351
acc cag gcg gcc agc ccc ctg gaa cgt cga
thr gln ala ala ser pro leu glu arg arg
1111/371
cga ttc cgc agc ttc ccg gtg cga cgg ctg
arg phe arg ser phe pro val arg arg leu
1171/391
ggg gtt cgg cga acg tta tct cag tgg aat
gly val arg arg thr leu ser gln trp asn
1231/411
gtt act gtt gaa agc cac acc cat gcc agt
val thr val glu ser his thr his ala ser
1291/431
ggc cta gta cag gaa gag caa cct agc gac
gly leu val gln glu glu gln pro ser asp
1351/451
ccg cag cag ccg gga acc cca ggt tat gct
pro gln gln pro gly thr pro gly tyr ala
1411/471
ttc gac tgg cgt tac cca ccg tcc ccg ccc
phe asp trp arg tyr pro pro ser pro pro
1471/491
gag gcg ttg ggt ggt acc cgg ccg ggt ctg
glu ala leu gly gly thr arg pro gly leu
1531/511
cct cct ggg atg gtt cgc caa cgc cct cgt
pro pro gly met val arg gln arg pro arg
1591/531
ata gcg gtg gtg tcc gcc ggc atc ggc ggc
ile ala val val ser ala gly ile gly gly
1651/551
gca ccc gcc ggc ccc agc ggc ggc cca gtg
ala pro ala gly pro ser gly gly pro val
1711/571
gca aac atg ccg ccg ggg tcg gtc gaa cag
ala asn met pro pro gly ser val glu gln
1771/591
atg ttg gaa acc gat ctg ggc cgc cag tcg
met leu glu thr asp leu gly arg gln ser
1831/611
gag ggg ctg atc ttg acc aac aac cac gtg
glu gly leu ile leu thr asn asn his val
1891/631
agt ccg ccg ccg aaa acg acg gta
ser pro pro pro lys thr thr val)

SEQ ID NOS:848-863 (continued 1)

FIG. 49A (continued 1)

1/1 ─SEQ ID NO: 864
cat cta cac cgc tca aca gcc ggg cca gac
(his leu his arg ser thr ala gly pro asp

31/11
gct gcc ggt cgg tgc tgc cga gaa ggc ggt
ala ala gly arg cys cys arg glu gly gly

61/21 SEQ ID NO: 865
gat ccg tgg cga gtt gtt cat gtc gcg gcg
asp pro trp arg val val his val ala ala

91/31
cac cac cgc cga cca acg ggt gct tgc cat
his his arg arg pro thr gly ala cys his

121/41
ccg tct gac caa cgg tag ttc gct gct gat
pro ser asp gln arg)AMB(phe ala ala asp

151/51
ctc caa aag tct caa gcc cac cga agc agt
leu gln lys ser gln ala his arg ser ser

181/61 SEQ ID NO: 866
cat gaa caa gct gcg ttg ggt gct att gat
his glu gln ala ala leu gly ala ile asp

211/71
cgt ggg tgg gat cgg ggt ggc ggt cgc cgc
arg gly trp asp arg gly gly gly arg arg

241/81
ggt ggc cgg ggg gat ggt cac ccg ggc cgg
gly gly arg gly asp gly his pro gly arg

271/91
gct gag gcc ggt ggg ccg cct cac cga agc
ala glu ala gly gly pro pro his arg ser

301/101
ggc cga gcg ggt ggc gcg aac cga cga cct
gly arg ala gly gly ala asn arg arg pro

331/111
gcg gcc cat ccc cgt ctt cgg cag cga cga
ala ala his pro arg leu arg gln arg arg

361/121
att ggc cag gct gac aga ggc att caa ttt
ile gly gln ala asp arg gly ile gln phe

391/131
aat gct gcg ggc gct ggc cga gtc acg gga
asn ala ala gly ala gly arg val thr gly

421/141
acg gca ggc aag gct ggt tac cga cgc cgg
thr ala gly lys ala gly tyr arg arg arg

451/151
aca tga att gcg tac ccc gct aac gtc gct
thr)OPA(ile ala tyr pro ala asn val ala
SEQ ID NO: 867

481/161
gcg cac caa tgt cga act ctt gat ggc ctc
ala his gln cys arg thr leu asp gly leu

511/171
gat ggc ccc ggg ggc tcc gcg gct acc caa
asp gly pro gly gly ser ala ala thr gln

541/181
gca gga gat ggt cga cct gcg tgc cga tgt
ala gly asp gly arg pro ala cys arg cys

571/191
gct ggc tca aat cga gga att gtc cac act
ala gly ser asn arg gly ile val his thr

601/201
ggt agg cga ttt ggt gga cct gtc ccg agg
gly arg arg phe gly gly pro val pro arg

631/211
cga cgc cgg aga agt ggt gca cga gcc ggt
arg arg arg arg ser gly ala arg ala gly

661/221
cga cat ggc tga cgt cgt cga ccg cag cct
arg his gly)OPA(arg arg arg pro gln pro

691/231
gga gcg ggt cag gcg gcg gcg caa cga tat
gly ala gly gln ala ala ala gln arg tyr

721/241 SEQ ID NO: 868
cct ttt cga cgt cga ggt gat tgg gtg gca
pro phe arg arg arg gly asp trp val ala

751/251
ggt tta tgg cga tac cgc tgg att gtc gcg
gly leu trp arg tyr arg trp ile val ala

781/261
gat ggc gct taa cct gat gga caa cgc cgc
asp gly ala)OCH(pro asp gly gln arg arg

811/271
gaa gtg gag ccc gcc ggg cgg cca cgt ggg
glu val glu pro ala gly arg pro arg gly

841/281 SEQ ID NO: 869
tgt cag gct gag cca gct cga cgc gtc gca
cys gln ala glu pro ala arg arg val ala

871/291
cgc tga gct ggt ggt ttc cga ccg cgg ccc
arg)OPA(ala gly gly phe arg pro arg pro
SEQ ID NO: 870

901/301
ggg cat tcc cgt gca gga gcg ccg tct ggt
gly his ser arg ala gly ala pro ser gly

931/311
gtt tga acg gtt tta ccg gtc ggc atc ggc
val)OPA(thr val leu pro val gly ile gly
SEQ ID NO: 871

961/321
acg ggc gtt gcc ggg ttc ggg cct cgg gtt
thr gly val ala gly phe gly pro arg val

991/331
ggc gat cgt caa aca ggt ggt gct caa cca
gly asp arg gln thr gly gly ala gln pro

1021/341
cgg cgg att gct gcg cat cga aga cac cga
arg arg ile ala ala his arg arg his arg

1051/351
ccc agg cgg cca gcc ccc tgg aac gtc gat
pro arg arg pro ala pro trp asn val asp

SEQ ID NOS:864-877

FIG. 49B

```
1081/361                                  1111/371
tta cgt gct gct ccc cgg ccg tcg gat gcc  gat tcc gca gct tcc cgg tgc gac ggc tgg
leu arg ala ala pro arg pro ser asp ala  asp ser ala ala ser arg cys asp gly trp
1141/381                                  1171/391
cgc tcg gag cac gga cat cga gaa ctc tcg  ggg ttc ggc gaa cgt tat ctc agt gga atc
arg ser glu his gly his arg glu leu ser  gly phe gly glu arg tyr leu ser gly ile
1201/401                                  1231/411
tca gtc cac gcg cgc aac cta gtt gtg cag  tta ctg ttg aaa gcc aca ccc atg cca gtc
ser val his ala arg asn leu val val gln  leu leu leu lys ala thr pro met pro val
1261/421                                  1291/431
cac gca tgg cca agt tgg ccc gag tag tgg  gcc tag tac agg aag agc aac cta gcg aca
his ala trp pro ser trp pro glu)AMB trp  ala AMB(tyr arg lys ser asn leu ala thr)
1321/441                                  1351/451       ‾—SEQ ID NO: 872
tga cga atc acc cac ggt att cgc cac cgc  cgc agc agc cgg gaa ccc cag gtt atg ctc
OPA(arg ile thr his gly ile arg his arg  arg ser ser arg glu pro gln val met leu
1381/461       ‾SEQ ID NO: 873            1411/471
agg ggc agc agc aaa cgt aca gcc agc agt  tcg act ggc gtt acc cac cgt ccc cgc ccc
arg gly ser ser lys arg thr ala ser ser  ser thr gly val thr his arg pro arg pro
1441/481                                  1471/491
cgc agc caa ccc agt acc gtc aac cct acg  agg cgt tgg gtg gta ccc ggc cgg gtc tga
arg ser gln pro ser thr val asn pro thr  arg arg trp val val pro gly arg val)OPA
1501/501                                  1531/511
tac ctg gcg tga ttc cga cca tga cgc ccc  ctc ctg gga tgg ttc gcc aac gcc ctc gtg
tyr leu ala OPA phe arg pro OPA(arg pro  leu leu gly trp phe ala asn ala leu val
1561/521       SEQ ID NO: 874 —           1591/531
cag gca tgt tgg cca tcg gcg cgg tga cga  tag cgg tgg tgt ccg ccg gca tcg gcg gcg
gln ala cys trp pro ser ala arg)OPA arg  AMB(arg trp cys pro pro ala ser ala ala
1621/541                                  1651/551       ‾SEQ ID NO: 875
cgg ccg cat ccc tgg tcg ggt tca acc ggg  cac ccg ccg gcc cca gcg gcg gcc cag tgg
arg pro his pro trp ser gly ser thr gly  his pro pro ala pro ala ala ala gln trp
1681/561                                  1711/571
ctg cca gcg cgg cgc caa gca tcc ccg cag  caa aca tgc cgc cgg ggt cgg tcg aac agg
leu pro ala arg arg gln ala ser pro gln  gln thr cys arg arg gly arg ser asn arg
1741/581                                  1771/591
tgg cgg cca agg tgg tgc cca gtg tcg tca  tgt tgg aaa ccg atc tgg gcc gcc agt cgg
trp arg pro arg trp cys pro val ser ser  cys trp lys pro ile trp ala ala ser arg
1801/601                                  1831/611
agg agg gct ccg gca tca ttc tgt ctg ccg  agg ggc tga tct tga cca aca acc acg tga
arg arg ala pro ala ser phe cys leu pro  arg gly)OPA ser OPA(pro thr thr thr)OPA
1861/621                                  1891/631       ‾SEQ ID NO: 876
tcg cgg cgg ccg cca agc ctc ccc tgg gca  gtc cgc cgc cga aaa cga cgg ta
(ser arg arg pro pro ser leu pro trp ala  val arg arg arg lys arg arg)
        SEQ ID NO: 877
```

SEQ ID NOS:864-877 (continued 1)

FIG. 49B (continued 1)

1/1 SEQ ID NO: 878
atc tac acc gct caa cag ccg ggc cag acg
ile tyr thr ala gln gln pro gly gln thr

31/11
ctg ccg gtc ggt gct gcc gag aag gcg gtg
leu pro val gly ala ala glu lys ala val

61/21 SEQ ID NO: 879
atc cgt ggc gag ttg ttc atg tcg cgg cgc
ile arg gly glu leu phe met ser arg arg

91/31
acc acc gcc gac caa cgg gtg ctt gcc atc
thr thr ala asp gln arg val leu ala ile

121/41
cgt ctg acc aac ggt agt tcg ctg ctg atc
arg leu thr asn gly ser ser leu leu ile

151/51
tcc aaa agt ctc aag ccc acc gaa gca gtc
ser lys ser leu lys pro thr glu ala val

181/61
atg aac aag ctg cgt tgg gtg cta ttg atc
met asn lys leu arg trp val leu leu ile

211/71
gtg ggt ggg atc ggg gtg gcg gtc gcc gcg
val gly gly ile gly val ala val ala ala

241/81
gtg gcc ggg ggg atg gtc acc cgg gcc ggg
val ala gly gly met val thr arg ala gly

271/91
ctg agg ccg gtg ggc cgc ctc acc gaa gcg
leu arg pro val gly arg leu thr glu ala

301/101
gcc gag cgg gtg gcg cga acc gac gac ctg
ala glu arg val ala arg thr asp asp leu

331/111
cgg ccc atc ccc gtc ttc ggc agc gac gaa
arg pro ile pro val phe gly ser asp glu

361/121
ttg gcc agg ctg aca gag gca ttc aat tta
leu ala arg leu thr glu ala phe asn leu

391/131
atg ctg cgg gcg ctg gcc gag tca cgg gaa
met leu arg ala leu ala glu ser arg glu

421/141
cgg cag gca agg ctg gtt acc gac gcc gga
arg gln ala arg leu val thr asp ala gly

451/151
cat gaa ttg cgt acc ccg cta acg tcg ctg
his glu leu arg thr pro leu thr ser leu

481/161
cgc acc aat gtc gaa ctc ttg atg gcc tcg
arg thr asn val glu leu leu met ala ser

511/171
atg gcc ccg ggg gct ccg cgg cta ccc aag
met ala pro gly ala pro arg leu pro lys

541/181
cag gag atg gtc gac ctg cgt gcc gat gtg
gln glu met val asp leu arg ala asp val

571/191
ctg gct caa atc gag gaa ttg tcc aca ctg
leu ala gln ile glu glu leu ser thr leu

601/201
gta ggc gat ttg gtg gac ctg tcc cga ggc
val gly asp leu val asp leu ser arg gly

631/211
gac gcc gga gaa gtg gtg cac gag ccg gtc
asp ala gly glu val val his glu pro val

661/221
gac atg gct gac gtc gtc gac cgc agc ctg
asp met ala asp val val asp arg ser leu

691/231
gag cgg gtc agg cgg cgg cgc aac gat atc
glu arg val arg arg arg arg asn asp ile

721/241
ctt ttc gac gtc gag gtg att ggg tgg cag
leu phe asp val glu val ile gly trp gln

751/251
gtt tat ggc gat acc gct gga ttg tcg cgg
val tyr gly asp thr ala gly leu ser arg

781/261
atg gcg ctt aac ctg atg gac aac gcc gcg
met ala leu asn leu met asp asn ala ala

811/271
aag tgg agc ccg ccg ggc ggc cac gtg ggt
lys trp ser pro pro gly gly his val gly

841/281
gtc agg ctg agc cag ctc gac gcg tcg cac
val arg leu ser gln leu asp ala ser his

871/291
gct gag ctg gtg gtt tcc gac cgc ggc ccg
ala glu leu val val ser asp arg gly pro

901/301
ggc att ccc gtg cag gag cgc cgt ctg gtg
gly ile pro val gln glu arg arg leu val

931/311
ttt gaa cgg ttt tac cgg tcg gca tcg gca
phe glu arg phe tyr arg ser ala ser ala

961/321
cgg gcg ttg ccg ggt tcg ggc ctc ggg ttg
arg ala leu pro gly ser gly leu gly leu

991/331
gcg atc gtc aaa cag gtg gtg ctc aac cac
ala ile val lys gln val val leu asn his

SEQ ID NOS:878-882

FIG. 49C

1021/341
ggc gga ttg ctg cgc atc gaa gac acc gac
gly gly leu leu arg ile glu asp thr asp
1081/361
tac gtg ctg ctc ccc ggc cgt cgg atg ccg
tyr val leu leu pro gly arg arg met pro
1141/381
gct cgg agc acg gac atc gag aac tct cgg
ala arg ser thr asp ile glu asn ser arg
1201/401
cag tcc acg cgc gca acc tag ttg tgc agt
gln ser thr arg ala thr)AMB(leu cys ser
1261/421   SEQ ID NO: 880
acg cat ggc caa gtt ggc ccg agt agt ggg
thr his gly gln val gly pro ser ser gly
1321/441
gac gaa tca ccc acg gta ttc gcc acc gcc
asp glu ser pro thr val phe ala thr ala
1381/461
ggg gca gca gca aac gta cag cca gca gtt
gly ala ala ala asn val gln pro ala val
1441/481
gca gcc aac cca gta ccg tca acc cta cga
ala ala asn pro val pro ser thr leu arg
1501/501
acc tgg cgt gat tcc gac cat gac gcc ccc
thr trp arg asp ser asp his asp ala pro
1561/521
agg cat gtt ggc cat cgg cgc ggt gac gat
arg his val gly his arg arg gly asp asp
1621/541
ggc cgc atc cct ggt cgg gtt caa ccg ggc
gly arg ile pro gly arg val gln pro gly
1681/561
tgc cag cgc ggc gcc aag cat ccc cgc agc
cys gln arg gly ala lys his pro arg ser
1741/581
ggc ggc caa ggt ggt gcc cag tgt cgt cat
gly gly gln gly gly ala gln cys arg his
1801/601
gga ggg ctc cgg cat cat tct gtc tgc cga
gly gly leu arg his his ser val cys arg
1861/621
cgc ggc ggc cgc caa gcc tcc cct ggg cag
arg gly gly arg gln ala ser pro gly gln

1051/351
cca ggc ggc cag ccc cct gga acg tcg att
pro gly gly gln pro pro gly thr ser ile
1111/371
att ccg cag ctt ccc ggt gcg acg gct ggc
ile pro gln leu pro gly ala thr ala gly
1171/391
ggt tcg gcg aac gtt atc tca gtg gaa tct
gly ser ala asn val ile ser val glu ser
1231/411
tac tgt tga aag cca cac cca tgc cag tcc
tyr cys)OPA(lys pro his pro cys gln ser
1291/431   SEQ ID NO: 881
cct agt aca gga aga gca acc tag cga cat
pro ser thr gly arg ala thr)AMB(arg his
1351/451   SEQ ID NO: 882
gca gca gcc ggg aac ccc agg tta tgc tca
ala ala ala gly asn pro arg leu cys ser
1411/471
cga ctg gcg tta ccc acc gtc ccc gcc ccc
arg leu ala leu pro thr val pro ala pro
1471/491
ggc gtt ggg tgg tac ccg gcc ggg tct gat
gly val gly trp tyr pro ala gly ser asp
1531/511
tcc tgg gat ggt tcg cca acg ccc tcg tgc
ser trp asp gly ser pro thr pro ser cys
1591/531
agc ggt ggt gtc cgc cgg cat cgg cgg cgc
ser gly gly val arg arg his arg arg arg
1651/551
acc cgc cgg ccc cag cgg cgg ccc agt ggc
thr arg arg pro gln arg arg pro ser gly
1711/571
aaa cat gcc gcc ggg gtc ggt cga aca ggt
lys his ala ala gly val gly arg thr gly
1771/591
gtt gga aac cga tct ggg ccg cca gtc gga
val gly asn arg ser gly pro pro val gly
1831/611
ggg gct gat ctt gac caa caa cca cgt gat
gly ala asp leu asp gln gln pro arg asp
1891/631
tcc gcc gcc gaa aac gac ggt a
ser ala ala glu asn asp gly)

SEQ ID NOS:878-882 (continued 1)

FIG. 49C (continued 1)

Coding sequence Rv0983 predicted by Cole et al., 1998 (Nature 393:537-544)
and containing seq60A:

1/1 ⟋ **SEQ ID NO: 883**   31/11
atg gcc aag ttg gcc cga gta gtg ggc cta gta cag gaa gag caa cct agc gac atg acg
(Met ala lys leu ala arg val val gly leu val gln glu glu gln pro ser asp met thr
61/21 ⟍ **SEQ ID NO: 884**   91/31
aat cac cca cgg tat tcg cca ccg ccg cag cag ccg gga acc cca ggt tat gct cag ggg
asn his pro arg tyr ser pro pro pro gln gln pro gly thr pro gly tyr ala gln gly
121/41   151/51
cag cag caa acg tac agc cag cag ttc gac tgg cgt tac cca ccg tcc ccg ccc ccg cag
gln gln gln thr tyr ser gln gln phe asp trp arg tyr pro pro ser pro pro pro gln
181/61   211/71
cca acc cag tac cgt caa ccc tac gag gcg ttg ggt ggt acc cgg ccg ggt ctg ata cct
pro thr gln tyr arg gln pro tyr glu ala leu gly gly thr arg pro gly leu ile pro
241/81   271/91
ggc gtg att ccg acc atg acg ccc cct cct ggg atg gtt cgc caa cgc cct cgt gca ggc
gly val ile pro thr met thr pro pro pro gly met val arg gln arg pro arg ala gly
301/101   331/111
atg ttg gcc atc ggc gcg gtg acg ata gcg gtg gtg tcc gcc ggc atc ggc ggc gcg gcc
met leu ala ile gly ala val thr ile ala val val ser ala gly ile gly gly ala ala
361/121   391/131
gca tcc ctg gtc ggg ttc aac cgg gca ccc gcc ggc ccc agc ggc ggc cca gtg gct gcc
ala ser leu val gly phe asn arg ala pro ala gly pro ser gly gly pro val ala ala
421/141   451/151
agc gcg gcg cca agc atc ccc gca gca aac atg ccg ccg ggg tcg gtc gaa cag gtg gcg
ser ala ala pro ser ile pro ala ala asn met pro pro gly ser val glu gln val ala
481/161   511/171
gcc aag gtg gtg ccc agt gtc gtc atg ttg gaa acc gat ctg ggc cgc cag tcg gag gag
ala lys val val pro ser val val met leu glu thr asp leu gly arg gln ser glu glu
541/181   571/191
ggc tcc ggc atc att ctg tct gcc gag ggg ctg atc ttg acc aac aac cac gtg atc gcg
gly ser gly ile ile leu ser ala glu gly leu ile leu thr asn asn his val ile ala
601/201   631/211
gcg gcc gcc aag cct ccc ctg ggc agt ccg ccg ccg aaa acg acg gta acc ttc tct gac
ala ala ala lys pro pro leu gly ser pro pro pro lys thr thr val thr phe ser asp
661/221   691/231
ggg cgg acc gca ccc ttc acg gtg gtg ggg gct gac ccc acc agt gat atc gcc gtc gtc
gly arg thr ala pro phe thr val val gly ala asp pro thr ser asp ile ala val val
721/241   751/251
cgt gtt cag ggc gtc tcc ggg ctc acc ccg atc tcc ctg ggt tcc tcc tcg gac ctg agg
arg val gln gly val ser gly leu thr pro ile ser leu gly ser ser ser asp leu arg
781/261   811/271
gtc ggt cag ccg gtg ctg gcg atc ggg tcg ccg ctc ggt ttg gag ggc acc gtg acc acg
val gly gln pro val leu ala ile gly ser pro leu gly leu glu gly thr val thr thr

SEQ ID NOS:883-884

FIG. 49D

841/281
ggg atc gtc agc gct ctc aac cgt cca gtg
gly ile val ser ala leu asn arg pro val
901/301
acc gtg ctg gac gcc att cag acc gac gcc
thr val leu asp ala ile gln thr asp ala
961/321
ctg gtg aac atg aac gct caa ctc gtc gga
leu val asn met asn ala gln leu val gly
1021/341
gac tca gcc gat gcg cag agc ggc tcg atc
asp ser ala asp ala gln ser gly ser ile
1081/361
gcc aag cgc atc gcc gac gag ttg atc agc
ala lys arg ile ala asp glu leu ile ser
1141/381
gtg cag gtg acc aat gac aaa gac acc ctg
val gln val thr asn asp lys asp thr leu
1201/401
ggt gct gcc gcg aac gct gga gtg ccg aag
gly ala ala ala asn ala gly val pro lys
1261/421
ccg atc aac agc gcg gac gcg ttg gtt gcc
pro ile asn ser ala asp ala leu val ala
1321/441
gtg gcg cta acc ttt cag gat ccc tcg ggc
val ala leu thr phe gln asp pro ser gly
1381/461
aag gcg gag cag tga
lys ala glu gln)OPA

871/291
tcg acg acc ggc gag gcc ggc aac cag aac
ser thr thr gly glu ala gly asn gln asn
931/311
gcg atc aac ccc ggt aac tcc ggg ggc gcg
ala ile asn pro gly asn ser gly gly ala
991/331
gtc aac tcg gcc att gcc acg ctg ggc gcg
val asn ser ala ile ala thr leu gly ala
1051/351
ggt ctc ggt ttt gcg att cca gtc gac cag
gly leu gly phe ala ile pro val asp gln
1111/371
acc ggc aag gcg tca cat gcc tcc ctg ggt
thr gly lys ala ser his ala ser leu gly
1171/391
ggc gcc aag atc gtc gaa gta gtg gcc ggt
gly ala lys ile val glu val val ala gly
1231/411
ggc gtc gtt gtc acc aag gtc gac gac cgc
gly val val val thr lys val asp asp arg
1291/431
gcc gtg cgg tcc aaa gcg ccg ggc gcc acg
ala val arg ser lys ala pro gly ala thr
1351/451
ggt agc cgc aca gtg caa gtc acc ctc ggc
gly ser arg thr val gln val thr leu gly

SEQ ID NOS:883-884 (continued 1)

FIG. 49D (continued 1)

ORF according to Cole et al., 1998 (Nature 393:537-544) and containing Rv0983

1/1 SEQ ID NO: 885   31/11
tga gcc agc tcg acg cgt cgc acg ctg agc tgg tgg ttt ccg acc gcg gcc cgg gca ttc
OPA(ala ser ser thr arg arg thr leu ser trp trp phe pro thr ala ala arg ala phe
61/21 SEQ ID NO: 886   91/31
ccg tgc agg agc gcc gtc tgg tgt ttg aac ggt ttt acc ggt cgg cat cgg cac ggg cgt
pro cys arg ser ala val trp cys leu asn gly phe thr gly arg his arg his gly arg
121/41   151/51
tgc cgg gtt cgg gcc tcg ggt tgg cga tcg tca aac agg tgg tgc tca acc acg gcg gat
cys arg val arg ala ser gly trp arg ser ser asn arg trp cys ser thr thr ala asp
181/61   211/71
tgc tgc gca tcg aag aca ccg acc cag gcg gcc agc ccc ctg gaa cgt cga ttt acg tgc
cys cys ala ser lys thr pro thr gln ala ala ser pro leu glu arg arg phe thr cys
241/81   271/91
tgc tcc ccg gcc gtc gga tgc cga ttc cgc agc ttc ccg gtg cga cgg ctg gcg ctc gga
cys ser pro ala val gly cys arg phe arg ser phe pro val arg arg leu ala leu gly
301/101   331/111
gca cgg aca tcg aga act ctc ggg gtt cgg cga acg tta tct cag tgg aat ctc agt cca
ala arg thr ser arg thr leu gly val arg arg thr leu ser gln trp asn leu ser pro
361/121   391/131
cgc gcg caa cct agt tgt gca gtt act gtt gaa agc cac acc cat gcc agt cca cgc atg
arg ala gln pro ser cys ala val thr val glu ser his thr his ala ser pro arg met
421/141   451/151
gcc aag ttg gcc cga gta gtg ggc cta gta cag gaa gag caa cct agc gac atg acg aat
ala lys leu ala arg val val gly leu val gln glu glu gln pro ser asp met thr asn
481/161   511/171
cac cca cgg tat tcg cca ccg ccg cag cag ccg gga acc cca ggt tat gct cag ggg cag
his pro arg tyr ser pro pro pro gln gln pro gly thr pro gly tyr ala gln gly gln
541/181   571/191
cag caa acg tac agc cag cag ttc gac tgg cgt tac cca ccg tcc ccg ccc ccg cag cca
gln gln thr tyr ser gln gln phe asp trp arg tyr pro pro ser pro pro pro gln pro
601/201   631/211
acc cag tac cgt caa ccc tac gag gcg ttg ggt ggt acc cgg ccg ggt ctg ata cct ggc
thr gln tyr arg gln pro tyr glu ala leu gly gly thr arg pro gly leu ile pro gly
661/221   691/231
gtg att ccg acc atg acg ccc cct cct ggg atg gtt cgc caa cgc cct cgt gca ggc atg
val ile pro thr met thr pro pro pro gly met val arg gln arg pro arg ala gly met
721/241   751/251
ttg gcc atc ggc gcg gtg acg ata gcg gtg gtg tcc gcc ggc atc ggc ggc gcg gcc gca
leu ala ile gly ala val thr ile ala val val ser ala gly ile gly gly ala ala ala
781/261   811/271
tcc ctg gtc ggg ttc aac cgg gca ccc gcc ggc ccc agc ggc ggc cca gtg gct gcc agc
ser leu val gly phe asn arg ala pro ala gly pro ser gly gly pro val ala ala ser
841/281   871/291
gcg gcg cca agc atc ccc gca gca aac atg ccg ccg ggg tcg gtc gaa cag gtg gcg gcc
ala ala pro ser ile pro ala ala asn met pro pro gly ser val glu gln val ala ala

SEQ ID NOS:885-886

FIG. 49E

```
901/301                                     931/311
aag gtg gtg ccc agt gtc gtc atg ttg gaa    acc gat ctg ggc cgc cag tcg gag gag ggc
lys val val pro ser val val met leu glu    thr asp leu gly arg gln ser glu glu gly
961/321                                     991/331
tcc ggc atc att ctg tct gcc gag ggg ctg    atc ttg acc aac aac cac gtg atc gcg gcg
ser gly ile ile leu ser ala glu gly leu    ile leu thr asn asn his val ile ala ala
1021/341                                    1051/351
gcc gcc aag cct ccc ctg ggc agt ccg ccg    ccg aaa acg acg gta acc ttc tct gac ggg
ala ala lys pro pro leu gly ser pro pro    pro lys thr thr val thr phe ser asp gly
1081/361                                    1111/371
cgg acc gca ccc ttc acg gtg gtg ggg gct    gac ccc acc agt gat atc gcc gtc gtc cgt
arg thr ala pro phe thr val val gly ala    asp pro thr ser asp ile ala val val arg
1141/381                                    1171/391
gtt cag ggc gtc tcc ggg ctc acc ccg atc    tcc ctg ggt tcc tcc tcg gac ctg agg gtc
val gln gly val ser gly leu thr pro ile    ser leu gly ser ser ser asp leu arg val
1201/401                                    1231/411
ggt cag ccg gtg ctg gcg atc ggg tcg ccg    ctc ggt ttg gag ggc acc gtg acc acg ggg
gly gln pro val leu ala ile gly ser pro    leu gly leu glu gly thr val thr thr gly
1261/421                                    1291/431
atc gtc agc gct ctc aac cgt cca gtg tcg    acg acc ggc gag gcc ggc aac cag aac acc
ile val ser ala leu asn arg pro val ser    thr thr gly glu ala gly asn gln asn thr
1321/441                                    1351/451
gtg ctg gac gcc att cag acc gac gcc gcg    atc aac ccc ggt aac tcc ggg ggc gcg ctg
val leu asp ala ile gln thr asp ala ala    ile asn pro gly asn ser gly gly ala leu
1381/461                                    1411/471
gtg aac atg aac gct caa ctc gtc gga gtc    aac tcg gcc att gcc acg ctg ggc gcg gac
val asn met asn ala gln leu val gly val    asn ser ala ile ala thr leu gly ala asp
1441/481                                    1471/491
tca gcc gat gcg cag agc ggc tcg atc ggt    ctc ggt ttt gcg att cca gtc gac cag gcc
ser ala asp ala gln ser gly ser ile gly    leu gly phe ala ile pro val asp gln ala
1501/501                                    1531/511
aag cgc atc gcc gac gag ttg atc agc acc    ggc aag gcg tca cat gcc tcc ctg ggt gtg
lys arg ile ala asp glu leu ile ser thr    gly lys ala ser his ala ser leu gly val
1561/521                                    1591/531
cag gtg acc aat gac aaa gac acc ctg ggc    gcc aag atc gtc gaa gta gtg gcc ggt ggt
gln val thr asn asp lys asp thr leu gly    ala lys ile val glu val val ala gly gly
1621/541                                    1651/551
gct gcc gcg aac gct gga gtg ccg aag ggc    gtc gtt gtc acc aag gtc gac gac cgc ccg
ala ala ala asn ala gly val pro lys gly    val val val thr lys val asp asp arg pro
1681/561                                    1711/571
atc aac agc gcg gac gcg ttg gtt gcc gcc    gtg cgg tcc aaa gcg ccg ggc gcc acg gtg
ile asn ser ala asp ala leu val ala ala    val arg ser lys ala pro gly ala thr val
1741/581                                    1771/591
gcg cta acc ttt cag gat ccc tcg ggc ggt    agc cgc aca gtg caa gtc acc ctc ggc aag
ala leu thr phe gln asp pro ser gly gly    ser arg thr val gln val thr leu gly lys
1801/601
gcg gag cag tga
ala glu gln)OPA
```

SEQ ID NOS:885-886 (continued 1)

FIG. 49E (continued 1)

Fragment amplified by PCR based on the sequence similarities with a serine
protease of the E.coli HtrA family (creation of an SnaBI site at the
3' end) and subcloned into the vector pJVEDa:

```
1/1    SEQ ID NO: 887                        31/11
gat ccg gcg ggg cgg gtg tcg gcg cag gcg tgg ctg gcg gtc acg gcg gtg cgg gcg gtg
(asp pro ala gly arg val ser ala gln ala trp leu ala val thr ala val arg ala val
61/21    SEQ ID NO: 888                      91/31
ccg ccg ggc tgt ggg gcg ccg gcg gcg gcg gtg gca atg gcg gga acg gcg ccg atg cca
pro pro gly cys gly ala pro ala ala ala val ala met ala gly thr ala pro met pro
121/41                                       151/51
aca tcg tca gcg gtg gag acg gtg gcc tcg gcg gtg ccg gtg gcg gtg gcg gat ggc tct
thr ser ser ala val glu thr val ala ser ala val pro val ala val ala asp gly ser
181/61                                       211/71
acg gcg acg gcg ggg ccg gcg gac acg gcg gac aag gcg caa tcg gcc tcg gcg gcg gcg
thr ala thr ala gly pro ala asp thr ala asp lys ala gln ser ala ser ala ala ala
241/81                                       271/91
ccg gcg gcg acg ggg gcc agg gcg gcg ccg gcc gcg gac tgt ggg gta ctg gcg gcg ccg
pro ala ala thr gly ala arg ala ala pro ala ala asp cys gly val leu ala ala pro
301/101                                      331/111
gcg gac acg gcg ggc aag gcg gtg gta ccg ggg gcc cac cgc tgc ccg gtc agg cag gca
ala asp thr ala gly lys ala val val pro gly ala his arg cys pro val arg gln ala
361/121                                      391/131
tgg gcg ccg cgg gtg gcg ccg gtg ggc tga tcg gca acg gcg ggg ccg gcg gcg acg gcg
trp ala pro arg val ala pro val gly)OPA(ser ala thr ala gly pro ala ala thr ala
421/141          SEQ ID NO: 889              451/151
gtg tcg gcg cgt ccg gcg ggg tcg ccg gag tag gcg gtg ccg gcg gga acg cca tgc tga
val ser ala arg pro ala gly ser pro glu)AMB(ala val pro ala gly thr pro cys)OPA
481/161                                      511/171    SEQ ID NO: 890
tcg ggc acg gcg gcg ccg gcg gcg ccg gcg gag aca gca gtt tcg cta atg gcg cgg ccg
(ser gly thr ala ala pro ala ala pro ala glu thr ala val ser leu met ala arg pro
541/181    SEQ ID NO: 891                    571/191
gcg gcg cgg gcg gtg ccg gag ggc acc tct tcg gca atg gcg ggt ccg gcg gcc acg gcg
ala ala arg ala val pro glu gly thr ser ser ala met ala gly pro ala ala thr ala
601/201                                      631/211
gag ccg tca cgg ccg gca aca ccg gta tcg gtg gcg ccg gcg gcg tcg gtg ggg acg cca
glu pro ser arg pro ala thr pro val ser val ala pro ala ala ser val gly thr pro
661/221                                      691/231
ggc tga tcg gcc acg gtg gcg ccg gcg gtg ccg gcg ggg acc gcg ccg gag cct tgg ttg
gly)OPA(ser ala thr val ala pro ala val pro ala gly thr ala pro glu pro trp leu
721/241    SEQ ID NO: 892                    751/251
gcc gtg acg gcg ggc ccg gtg gga acg ggg gcg ctg gcg gcc agc tat acg gca acg gcg
ala val thr ala gly pro val gly thr gly ala leu ala ala ser tyr thr ala thr ala
781/261                                      811/271
gcg acg gcg ccc ccg gca ccg gcg gaa cac tgc agg cgg cgg tga gcg gat tgg tga cgg
ala thr ala pro pro ala pro ala glu his cys arg arg arg)OPA ala asp trp OPA(arg
841/281                                      871/291        SEQ ID NO: 893)
ctt tgt tcg gtg cac ccg gcc aac ccg gcg aca ccg gcc aac ccg gct agc ccc gat caa
leu cys ser val his pro ala asn pro ala thr pro ala asn pro ala ser pro asp gln
901/301                                      931/311
cga ggg ttt cgg tgc cgg tcc ggg gca tgg cca tcc gct gag ctg gcg atc tgg act acg
arg gly phe arg cys arg ser gly ala trp pro ser ala glu leu ala ile trp thr thr
961/321                                      991/331
ttg gtg tag aaa aat cct gcc gcc cgg acc ctt aag gct ggg aca att tct gat agc tac
leu val)AMB(lys asn pro ala ala arg thr leu lys ala gly thr ile ser asp ser tyr
1021/341    SEQ ID NO: 894                   1051/351
ccc gac aca gga ggt tac ggg atg agc aat tcg cgc cgc cgc tca ctc agg tgg tca tgg
pro asp thr gly gly tyr gly met ser asn ser arg arg arg ser leu arg trp ser trp
1081/361                                     1111/371
ttg ctg agc gtg ctg gct gcc gtc ggg ctg ggc ctg gcc acg gcg ccg gcc cag gcg gcc
leu leu ser val leu ala ala val gly leu gly leu ala thr ala pro ala gln ala ala
1141/381
ccg ccg gcc ttg tcg cag gac cgg tt
pro pro ala leu ser gln asp arg)
```

SEQ ID NOS:887-894

FIG. 50A

1/1 ─ SEQ ID NO: 895   31/11
atc cgg cgg ggc ggg tgt cgg cgc agg cgt ggc tgg cgg tca cgg cgg tgc ggg cgg tgc
(ile arg arg gly gly cys arg arg arg cgt gly trp arg ser arg arg cys gly arg cys

61/21 ─ SEQ ID NO: 896   91/31
cgc cgg gct gtg ggg cgc cgg cgg cgg cgg tgg caa tgg cgg gaa cgg cgc cga tgc caa
arg arg ala val gly arg arg arg arg arg trp gln trp arg glu arg arg arg cys gln

121/41   151/51
cat cgt cag cgg tgg aga cgg tgg cct cgg cgg tgc cgg tgg cgg tgg cgg atg gct cta
his arg gln arg trp arg arg trp pro arg arg cys arg trp arg trp arg met ala leu

181/61   211/71
cgg cga cgg cgg ggc cgg cgg aca cgg cgg aca agg cgc aat cgg cct cgg cgg cgg cgc
arg arg arg arg gly arg arg thr arg arg thr arg arg asn arg pro arg arg arg arg

241/81   271/91
cgg cgg cga cgg ggg cca ggg cgg cgc cgg ccg cgg act gtg ggg tac tgg cgg cgc cgg
arg arg arg arg gly pro gly arg arg arg pro arg thr val gly tyr trp arg arg arg

301/101   331/111
cgg aca cgg cgg gca agg cgg tgg tac cgg ggg ccc acc gct gcc cgg tca ggc agg cat
arg thr arg arg ala arg arg trp tyr arg gly pro thr ala ala arg ser gly arg his

361/121   391/131
ggg cgc cgc ggg tgg cgc cgg tgg gct gat cgg caa cgg cgg ggc cgg cgg cga cgg cgg
gly arg arg gly trp arg arg trp ala asp arg gln arg arg gly arg arg arg arg arg

421/141   451/151
tgt cgg cgc gtc cgg cgg ggt cgc cgg agt agg cgg tgc cgg cgg gaa cgc cat gct gat
cys arg arg val arg arg gly arg arg ser arg arg cys arg arg glu arg his ala asp

481/161   511/171
cgg gca cgg cgg cgc cgg cgg cgc cgg cgg aga cag cag ttt cgc taa tgg cgc ggc cgg
arg ala arg arg arg arg arg arg arg arg arg gln gln phe arg)OCH(trp arg gly arg

541/181   571/191   SEQ ID NO: 897
cgg cgc ggg cgg tgc cgg agg gca cct ctt cgg caa tgg cgg gtc cgg cgg cca cgg cgg
arg arg gly arg cys arg arg ala pro leu arg gln trp arg val arg arg pro arg arg

601/201   631/211
agc cgt cac ggc cgg caa cac cgg tat cgg tgg cgc cgg cgg cgt cgg tgg gga cgc cag
ser arg his gly arg gln his arg tyr arg trp arg arg arg arg trp gly arg gln

661/221   691/231
gct gat cgg cca cgg tgg cgc cgg cgg tgc cgg cgg gga ccg cgc cgg agc ctt ggt tgg
ala asp arg pro arg trp arg arg arg cys arg arg gly pro arg arg ser leu gly trp

721/241   751/251
ccg tga cgg cgg gcc cgg tgg gaa cgg ggg cgc tgg cgg cca gct ata cgg caa cgg cgg
pro)OPA(arg arg ala arg trp glu arg gly arg trp arg pro ala ile arg gln arg arg

781/261 ─── SEQ ID NO: 898   811/271
cga cgg cgc ccc cgg cac cgg cgg aac act gca ggc ggc ggt gag cgg att ggt gac ggc
arg arg arg pro arg his arg arg asn thr ala gly gly gly glu arg ile gly asp gly

841/281   871/291
ttt gtt cgg tgc acc cgg cca acc cgg cga cac cgg cca acc cgg cta gcc ccg atc aac
phe val arg cys thr arg pro thr arg arg his arg pro thr arg leu ala pro ile asn

901/301   931/311
gag ggt ttc ggt gcc ggt ccg ggg cat ggc cat ccg ctg agc tgg cga tct gga cta cgt
glu gly phe gly ala gly pro gly his gly his pro leu ser trp arg ser gly leu arg

961/321   991/331
tgg tgt aga aaa atc ctg ccg ccc gga ccc tta agg ctg gga caa ttt ctg ata gct acc
trp cys arg lys ile leu pro pro gly pro leu arg leu gly gln phe leu ile ala thr

1021/341   1051/351
ccg aca cag gag gtt acg gga tga gca att cgc gcc gcc gct cac tca ggt ggt cat ggt
pro thr gln glu val thr gly)OPA(ala ile arg ala ala ala his ser gly gly his gly

1081/361   SEQ ID NO: 899 ──   1111/371
tgc tga gcg tgc tgg ctg ccg tcg ggc tgg gcc tgg cca cgg cgc cgg ccc agg cgg ccc
cys)OPA(ala cys trp leu pro ser gly trp ala trp pro arg arg arg pro arg arg pro

1141/381 ─ SEQ ID NO: 900
cgc cgg cct tgt cgc agg acc ggt t
arg arg pro cys arg arg thr gly)

SEQ ID NOS:895-900

FIG.50B

1/1 SEQ ID NO: 901   31/11
tcc ggc ggg gcg ggt gtc ggc gca ggc gtg gct ggc ggt cac ggc ggt gcg ggc ggt gcc
(ser gly gly ala gly val gly ala gly val ala gly gly his gly gly ala gly gly ala
61/21 SEQ ID NO: 902   91/31
gcc ggg ctg tgg ggc gcc ggc ggc ggc ggt ggc aat ggc ggg aac ggc gcc gat gcc aac
ala gly leu trp gly ala gly gly gly gly gly asn gly gly asn gly ala asp ala asn
121/41   151/51
atc gtc agc ggt gga gac ggt ggc ctc ggc ggt gcc ggt ggc ggt ggc gga tgg ctc tac
ile val ser gly gly asp gly gly leu gly gly ala gly gly gly gly gly trp leu tyr
181/61   211/71
ggc gac ggc ggg gcc ggc gga cac ggc gga caa ggc gca atc ggc ctc ggc ggc ggc gcc
gly asp gly gly ala gly gly his gly gly gln gly ala ile gly leu gly gly gly ala
241/81   271/91
ggc ggc gac ggg ggc cag ggc ggc gcc ggc cgc gga ctg tgg ggt act ggc ggc gcc ggc
gly gly asp gly gly gln gly gly ala gly arg gly leu trp gly thr gly gly ala gly
301/101   331/111
gga cac ggc ggg caa ggc ggt ggt acc ggg ggc cca ccg ctg ccc ggt cag gca ggc atg
gly his gly gly gln gly gly gly thr gly gly pro pro leu pro gly gln ala gly met
361/121   391/131
ggc gcc gcg ggt ggc gcc ggt ggg ctg atc ggc aac ggc ggg gcc ggc ggc gac ggc ggt
gly ala ala gly gly ala gly gly leu ile gly asn gly gly ala gly gly asp gly gly
421/141   451/151
gtc ggc gcg tcc ggc ggg gtc gcc gga gta ggc ggt gcc ggc ggg aac gcc atg ctg atc
val gly ala ser gly gly val ala gly val gly gly ala gly gly asn ala met leu ile
481/161   511/171
ggg cac ggc ggc gcc ggc ggc gcc ggc gga gac agc agt ttc gct aat ggc gcg gcc ggc
gly his gly gly ala gly gly ala gly gly asp ser ser phe ala asn gly ala ala gly
541/181   571/191
ggc gcg ggc ggt gcc gga ggg cac ctc ttc ggc aat ggc ggg tcc ggc ggc cac ggc gga
gly ala gly gly ala gly gly his leu phe gly asn gly gly ser gly gly his gly gly
601/201   631/211
gcc gtc acg gcc ggc aac acc ggt atc ggt ggc gcc ggc ggc gtc ggt ggg gac gcc agg
ala val thr ala gly asn thr gly ile gly gly ala gly gly val gly gly asp ala arg
661/221   691/231
ctg atc ggc cac ggt ggc gcc ggc ggt gcc ggc ggg gac cgc gcc gga gcc ttg gtt ggc
leu ile gly his gly gly ala gly gly ala gly gly asp arg ala gly ala leu val gly
721/241   751/251
cgt gac ggc ggg ccc ggt ggg aac ggg ggc gct ggc ggc cag cta tac ggc aac ggc ggc
arg asp gly gly pro gly gly asn gly gly ala gly gly gln leu tyr gly asn gly gly
781/261   811/271
gac ggc gcc ccc ggc acc ggc gga aca ctg cag gcg gcg gtg agc gga ttg gtg acg gct
asp gly ala pro gly thr gly gly thr leu gln ala ala val ser gly leu val thr ala
841/281   871/291
ttg ttc ggt gca ccc ggc caa ccc ggc gac acc ggc caa ccc ggc tag ccc cga tca acg
leu phe gly ala pro gly gln pro gly asp thr gly gln pro gly)AMB(pro arg ser thr
901/301   931/311   SEQ ID NO: 903
agg gtt tcg gtg ccg gtc cgg ggc atg gcc atc cgc tga gct ggc gat ctg gac tac gtt
arg val ser val pro val arg gly met ala ile arg)OPA(ala gly asp leu asp tyr val
961/321   991/331   SEQ ID NO: 904
ggt gta gaa aaa tcc tgc cgc ccg gac cct taa ggc tgg gac aat ttc tga tag cta ccc
gly val glu lys ser cys arg pro asp pro)OCH(gly trp asp asn phe)OPA AMB(leu pro
1021/341   SEQ ID NO: 905   1051/351   SEQ ID NO 906
cga cac agg agg tta cgg gat gag caa ttc gcg ccg ccg ctc act cag gtg gtc atg gtt
arg his arg arg leu arg asp glu gln phe ala pro pro leu thr gln val val met val
1081/361   1111/371
gct gag cgt gct ggc tgc cgt cgg gct ggg cct ggc cac ggc gcc ggc cca ggc ggc ccc
ala glu arg ala gly cys arg arg ala gly pro gly his gly ala gly pro gly gly pro
1141/381
gcc ggc ctt gtc gca gga ccg gtt
ala gly leu val ala gly pro val)

SEQ ID NOS:901-906

FIG.50C

Coding sequence Rv0125 predicted by Cole et al., 1998 (Nature 393:537-544)
and containing seq50A:

1/1 ─SEQ ID NO: 907   31/11
atg agc aat tcg cgc cgc cgc tca ctc agg tgg tca tgg ttg ctg agc gtg ctg gct gcc
(Met ser asn ser arg arg arg ser leu arg trp ser trp leu leu ser val leu ala ala
61/21 ─SEQ ID NO: 908   91/31
gtc ggg ctg ggc ctg gcc acg gcg ccg gcc cag gcg gcc ccg ccg gcc ttg tcg cag gac
val gly leu gly leu ala thr ala pro ala gln ala ala pro pro ala leu ser gln asp
121/41   151/51
cgg ttc gcc gac ttc ccc gcg ctg ccc ctc gac ccg tcc gcg atg gtc gcc caa gtg ggg
arg phe ala asp phe pro ala leu pro leu asp pro ser ala met val ala gln val gly
181/61   211/71
cca cag gtg gtc aac atc aac acc aaa ctg ggc tac aac aac gcc gtg ggc gcc ggg acc
pro gln val val asn ile asn thr lys leu gly tyr asn asn ala val gly ala gly thr
241/81   271/91
ggc atc gtc atc gat ccc aac ggt gtc gtg ctg acc aac aac cac gtg atc gcg ggc gcc
gly ile val ile asp pro asn gly val val leu thr asn asn his val ile ala gly ala
301/101   331/111
acc gac atc aat gcg ttc agc gtc ggc tcc ggc caa acc tac ggc gtc gat gtg gtc ggg
thr asp ile asn ala phe ser val gly ser gly gln thr tyr gly val asp val val gly
361/121   391/131
tat gac cgc acc cag gat gtc gcg gtg ctg cag ctg cgc ggt gcc ggt ggc ctg ccg tcg
tyr asp arg thr gln asp val ala val leu gln leu arg gly ala gly gly leu pro ser
421/141   451/151
gcg gcg atc ggt ggc ggc gtc gcg gtt ggt gag ccc gtc gtc gcg atg ggc aac agc ggt
ala ala ile gly gly gly val ala val gly glu pro val val ala met gly asn ser gly
481/161   511/171
ggg cag ggc gga acg ccc cgt gcg gtg cct ggc agg gtg gtc gcg ctc ggc caa acc gtg
gly gln gly gly thr pro arg ala val pro gly arg val val ala leu gly gln thr val
541/181   571/191
cag gcg tcg gat tcg ctg acc ggt gcc gaa gag aca ttg aac ggg ttg atc cag ttc gat
gln ala ser asp ser leu thr gly ala glu glu thr leu asn gly leu ile gln phe asp
601/201   631/211
gcc gcg atc cag ccc ggt gat tcg ggc ggg ccc gtc gtc aac ggc cta gga cag gtg gtc
ala ala ile gln pro gly asp ser gly gly pro val val asn gly leu gly gln val val
661/221   691/231
ggt atg aac acg gcc gcg tcc gat aac ttc cag ctg tcc cag ggt ggg cag gga ttc gcc
gly met asn thr ala ala ser asp asn phe gln leu ser gln gly gly gln gly phe ala
721/241   751/251
att ccg atc ggg cag gcg atg gcg atc gcg ggc cag atc cga tcg ggt ggg ggg tca ccc
ile pro ile gly gln ala met ala ile ala gly gln ile arg ser gly gly gly ser pro
781/261   811/271
acc gtt cat atc ggg cct acc gcc ttc ctc ggc ttg ggt gtt gtc gac aac aac ggc aac
thr val his ile gly pro thr ala phe leu gly leu gly val val asp asn asn gly asn
841/281   871/291
ggc gca cga gtc caa cgc gtg gtc ggg agc gct ccg gcg gca agt ctc ggc atc tcc acc
gly ala arg val gln arg val val gly ser ala pro ala ala ser leu gly ile ser thr
901/301   931/311
ggc gac gtg atc acc gcg gtc gac ggc gct ccg atc aac tcg gcc acc gcg atg gcg gac
gly asp val ile thr ala val asp gly ala pro ile asn ser ala thr ala met ala asp
961/321   991/331
gcg ctt aac ggg cat cat ccc ggt gac gtc atc tcg gtg acc tgg caa acc aag tcg ggc
ala leu asn gly his his pro gly asp val ile ser val thr trp gln thr lys ser gly
1021/341   1051/351
ggc acg cgt aca ggg aac gtg aca ttg gcc gag gga ccc ccg gcc tga
gly thr arg thr gly asn val thr leu ala glu gly pro pro ala)OPA

SEQ ID NOS:907-908

FIG.50D

ORF according to Cole et al., 1998 (Nature 393:537-544) and containing Rv0125:

```
1/1 ___SEQ ID NO: 909                    31/11
tag aaa aat cct gcc gcc cgg acc ctt aag gct ggg aca att tct gat agc tac ccc gac
AMB(lys asn pro ala ala arg thr leu lys ala gly thr ile ser asp ser tyr pro asp
61/21 ___SEQ ID NO: 910                  91/31
aca gga ggt tac ggg atg agc aat tcg cgc cgc cgc tca ctc agg tgg tca tgg ttg ctg
thr gly gly tyr gly met ser asn ser arg arg arg ser leu arg trp ser trp leu leu
121/41                                   151/51
agc gtg ctg gct gcc gtc ggg ctg ggc ctg gcc acg gcg ccg gcc cag gcg gcc ccg ccg
ser val leu ala ala val gly leu gly leu ala thr ala pro ala gln ala ala pro pro
181/61                                   211/71
gcc ttg tcg cag gac cgg ttc gcc gac ttc ccc gcg ctg ccc ctc gac ccg tcc gcg atg
ala leu ser gln asp arg phe ala asp phe pro ala leu pro leu asp pro ser ala met
241/81                                   271/91
gtc gcc caa gtg ggg cca cag gtg gtc aac atc aac acc aaa ctg ggc tac aac aac gcc
val ala gln val gly pro gln val val asn ile asn thr lys leu gly tyr asn asn ala
301/101                                  331/111
gtg ggc gcc ggg acc ggc atc gtc atc gat ccc aac ggt gtc gtg ctg acc aac aac cac
val gly ala gly thr gly ile val ile asp pro asn gly val val leu thr asn asn his
361/121                                  391/131
gtg atc gcg ggc gcc acc gac atc aat gcg ttc agc gtc ggc tcc ggc caa acc tac ggc
val ile ala gly ala thr asp ile asn ala phe ser val gly ser gly gln thr tyr gly
421/141                                  451/151
gtc gat gtg gtc ggg tat gac cgc acc cag gat gtc gcg gtg ctg cag ctg cgc ggt gcc
val asp val val gly tyr asp arg thr gln asp val ala val leu gln leu arg gly ala
481/161                                  511/171
ggt ggc ctg ccg tcg gcg gcg atc ggt ggc ggc gtc gcg gtt ggt gag ccc gtc gtc gcg
gly gly leu pro ser ala ala ile gly gly gly val ala val gly glu pro val val ala
541/181                                  571/191
atg ggc aac agc ggt ggg cag ggc gga acg ccc cgt gcg gtg cct ggc agg gtg gtc gcg
met gly asn ser gly gly gln gly gly thr pro arg ala val pro gly arg val val ala
601/201                                  631/211
ctc ggc caa acc gtg cag gcg tcg gat tcg ctg acc ggt gcc gaa gag aca ttg aac ggg
leu gly gln thr val gln ala ser asp ser leu thr gly ala glu glu thr leu asn gly
661/221                                  691/231
ttg atc cag ttc gat gcc gcg atc cag ccc ggt gat tcg ggc ggg ccc gtc gtc aac ggc
leu ile gln phe asp ala ala ile gln pro gly asp ser gly gly pro val val asn gly
721/241                                  751/251
cta gga cag gtg gtc ggt atg aac acg gcc gcg tcc gat aac ttc cag ctg tcc cag ggt
leu gly gln val val gly met asn thr ala ala ser asp asn phe gln leu ser gln gly
781/261                                  811/271
ggg cag gga ttc gcc att ccg atc ggg cag gcg atg gcg atc gcg ggc cag atc cga tcg
gly gln gly phe ala ile pro ile gly gln ala met ala ile ala gly gln ile arg ser
841/281                                  871/291
ggt ggg ggg tca ccc acc gtt cat atc ggg cct acc gcc ttc ctc ggc ttg ggt gtt gtc
gly gly gly ser pro thr val his ile gly pro thr ala phe leu gly leu gly val val
901/301                                  931/311
gac aac aac ggc aac ggc gca cga gtc caa cgc gtg gtc ggg agc gct ccg gcg gca agt
asp asn asn gly asn gly ala arg val gln arg val val gly ser ala pro ala ala ser
961/321                                  991/331
ctc ggc atc tcc acc ggc gac gtg atc acc gcg gtc gac ggc gct ccg atc aac tcg gcc
leu gly ile ser thr gly asp val ile thr ala val asp gly ala pro ile asn ser ala
1021/341                                 1051/351
acc gcg atg gcg gac gcg ctt aac ggg cat.cat ccc ggt gac gtc atc tcg gtg acc tgg
thr ala met ala asp ala leu asn gly his his pro gly asp val ile ser val thr trp
1081/361                                 1111/371
caa acc aag tcg ggc ggc acg cgt aca ggg aac gtg aca ttg gcc gag gga ccc ccg gcc
gln thr lys ser gly gly thr arg thr gly asn val thr leu ala glu gly pro pro ala)
1141/381
tga
OPA
```
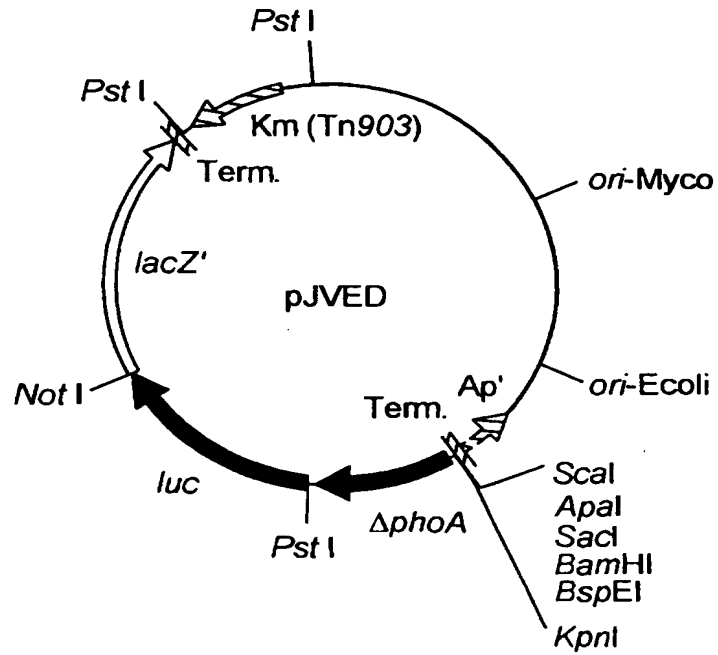
SEQ ID NOS:909-910

FIG.50E

FIG. 51A

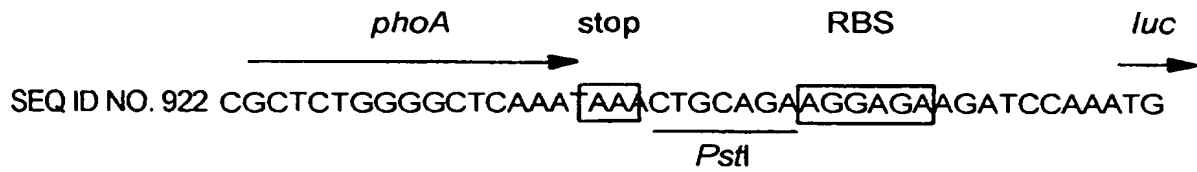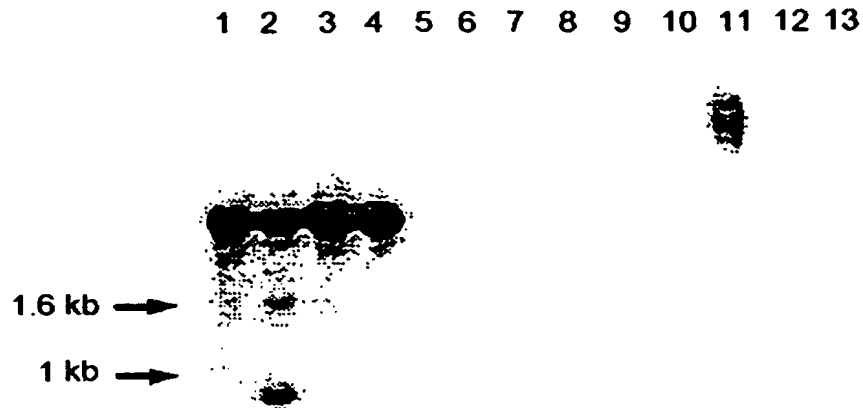**FIG. 51**

Experiment of molecular hybridization of a specific to DP428 on the genomic DNA of various mycobacterial species

1: *M. tuberculosis* 2: *M. bovis* 3: BCG 4: *M. africanum* 5: cancelled 6: *M. fortuitum* 7: *M. simiae* 8: *M. avium* 9: *M. chelonae* 10: *M. flavescens* 11: *M. gordonae* 12: *M. marinum* 13: *M. kansasii*
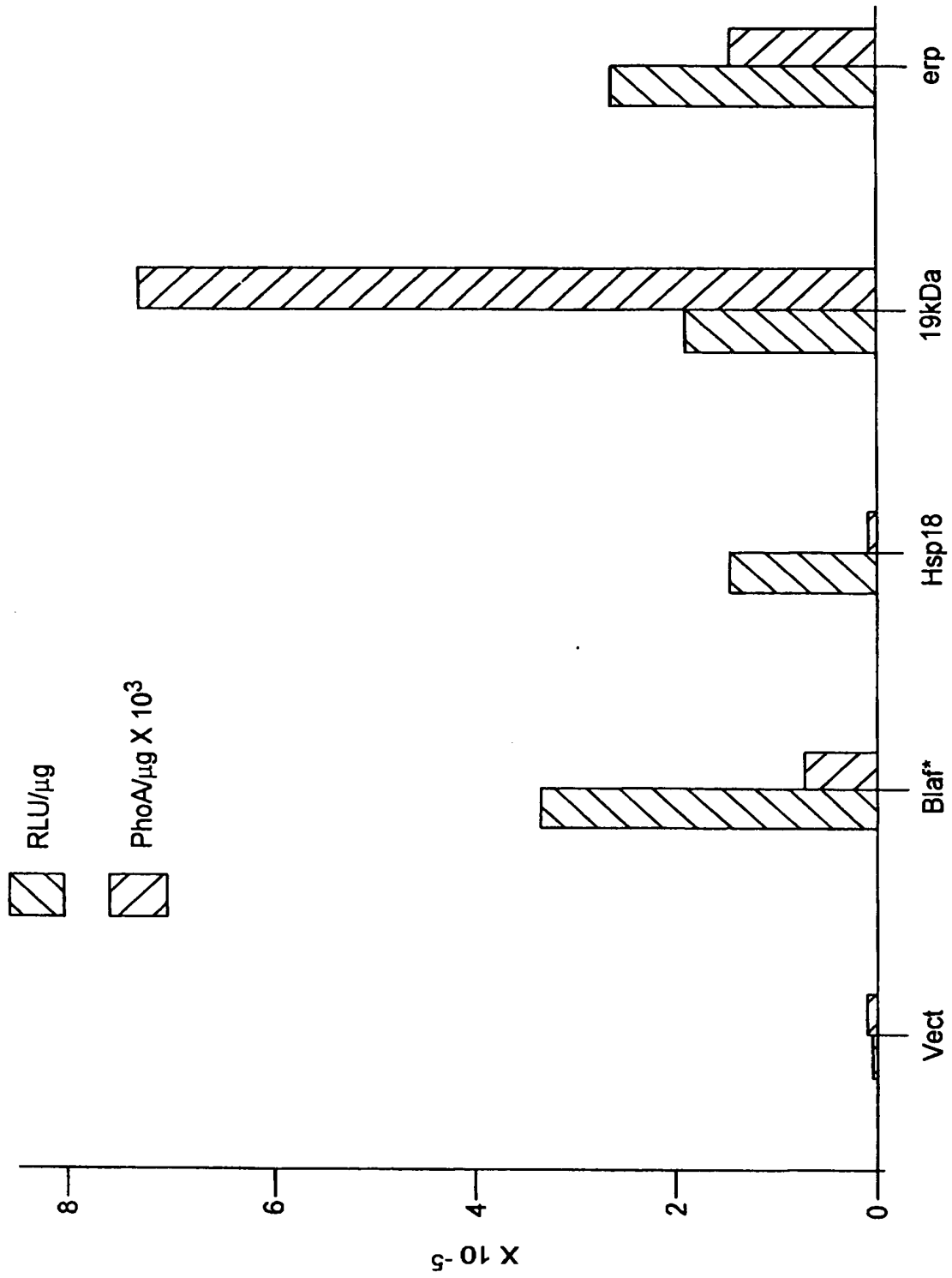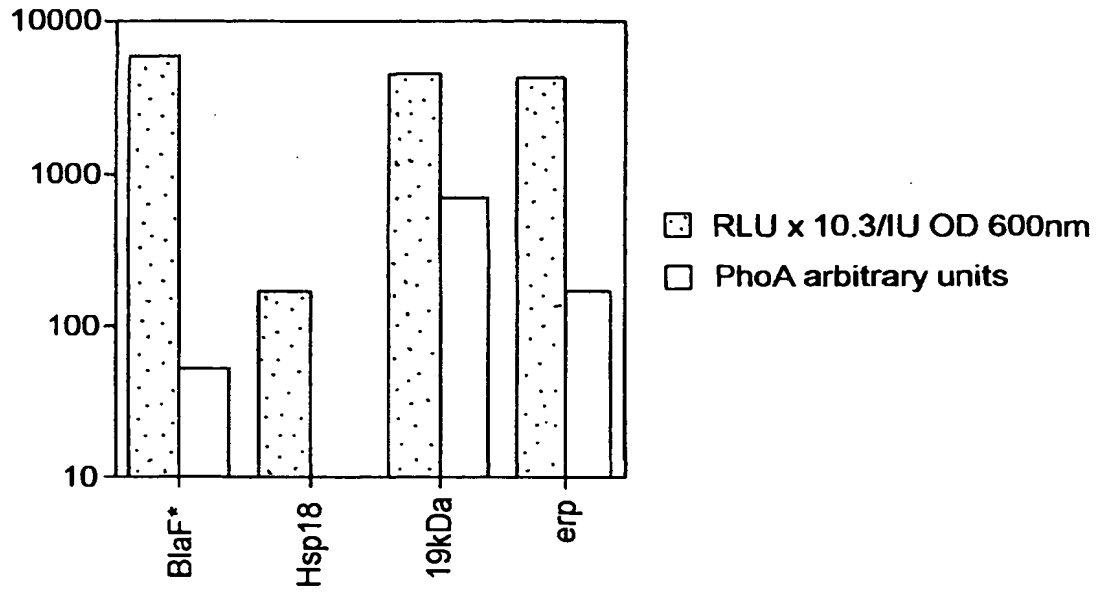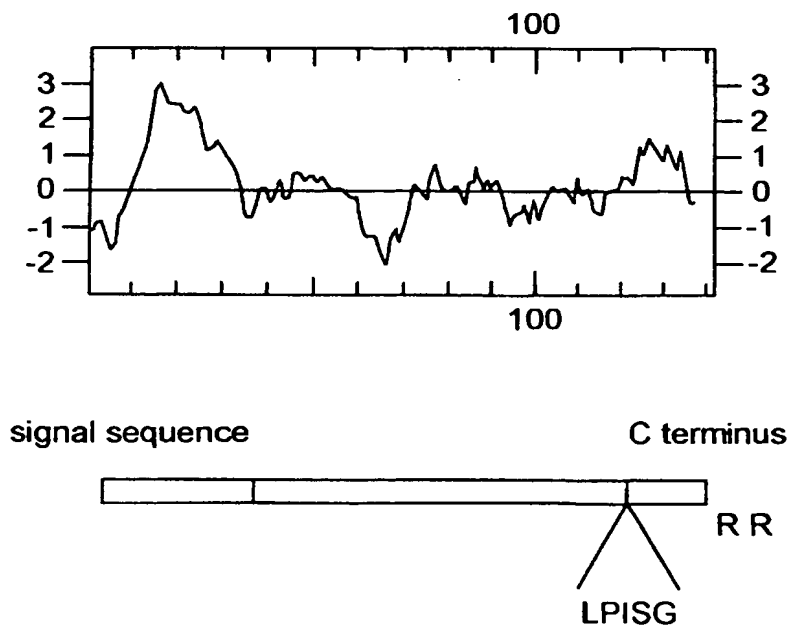
FIG. 52

FIG. 53

FIG. 54
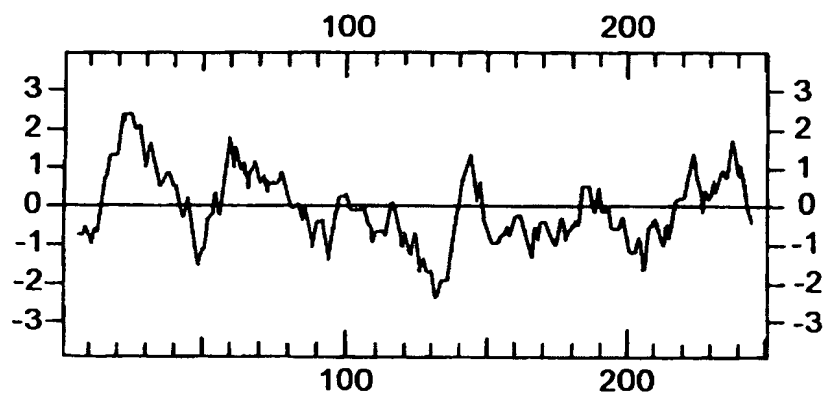

signal sequence   C terminus

FIG. 55

FIG. 56

FIG. 57A

FIG. 57B